Institute for International Economic Studies
Monograph Series No. 107

Stockholm University

Essays in Macroeconomics and Household Finance

Richard Foltyn

Doctoral Thesis in Economics at Stockholm University, Sweden 2020

Essays in Macroeconomics and Household Finance

Richard Foltyn

Academic dissertation for the Degree of Doctor of Philosophy in Economics at Stockholm University to be publicly defended on Friday 5 June 2020 at 14.00 in Nordenskiöldsalen, Geovetenskapens hus, Svante Arrhenius väg 12.

Abstract

Experience-based Learning, Stock Market Participation and Portfolio Choice

Recent evidence suggests that lifetime experiences play an important role in determining households' investment choices. I incorporate these findings and the fact that household portfolios are underdiversified into an otherwise standard life-cycle model and examine to what extent they can help resolve long-standing puzzles in the literature regarding stock market participation and the fraction of financial wealth invested in risky assets. I show that experience-based learning about returns creates a positive correlation between a household's position in the wealth distribution and its optimism about future returns. The wealthy consequently increase their investment in risky assets, while participation is limited among poor households. I find that in a reasonably calibrated quantitative model, this mechanism is able to close approximately half of the gap between the participation rates observed in the data and the predictions from standard models.

Health Dynamics and Heterogeneous Life Expectancies

In this paper, we provide improved estimates for age-dependent health transitions and survival probabilities for different subsamples of the US population. The estimated yearly transition matrices can be used in any life-cycle model where health and survival dynamics are of interest. The results show substantial heterogeneity in life expectancy in the population. For a 70-year-old man in excellent health, the probability of reaching his 80th birthday is around 75%, while the corresponding probability for a man in poor health is just below 40%.

Subjective Life Expectancies, Time Preference Heterogeneity and Wealth Inequality

Time preference heterogeneity is one of the potential sources of wealth inequality, but preferences are difficult to measure and quantify. In this paper, we investigate one source of time preference heterogeneity: heterogeneity in life expectancy. We document a systematic bias in subjective survival beliefs within cohorts that exacerbates the heterogeneity found in the population: individuals with a low survival probability relative to their peers underestimate their life expectancies, while individuals with a high survival probability overestimate theirs. We introduce survival heterogeneity into an otherwise standard overlapping-generations model and let survival probabilities and beliefs evolve stochastically according to a health and death process estimated from micro data. We find strong effects of life expectancy heterogeneity on within-cohort wealth inequality, but small effects on economy-wide wealth inequality.

On the Redistributive Effects of Government Bailouts in the Mortgage Market

In this paper we study the aggregate and distributional consequences of government bailout guarantees in the US mortgage market. We construct a model with aggregate risk in which competitive financial intermediaries issue mortgages to households that can default on their debt. Default probabilities are priced into mortgage interest rates unless a government bailout guarantee makes the lenders whole even in economic downturns in which foreclosure rates surge. We use the model to assess the extent to which bailout guarantees lead to excessive mortgage lending, household leverage and foreclosures in episodes of housing crises, as well as excess volatility in house prices in severe recessions.

Stockholm 2020
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-180800

ISBN 978-91-7911-158-8
ISBN 978-91-7911-159-5
ISSN 0346-6892

Stockholm University

Department of Economics

Stockholm University, 106 91 Stockholm

ESSAYS IN MACROECONOMICS AND HOUSEHOLD FINANCE

Richard Foltyn

Stockholm University

Essays in Macroeconomics and Household Finance

Richard Foltyn

ISBN print 978-91-7911-158-8
ISBN PDF 978-91-7911-159-5
ISSN 0346-6892

Cover photo: Heterogeneous sheep on a unit interval (by Evgenii Zotov, Getty Images)
Back-cover photo by Michaela de Verdier

Printed in Sweden by Universitetsservice US-AB, Stockholm 2020

Doctoral dissertation
Department of Economics
Stockholm University

Abstracts

Experience-based Learning, Stock Market Participation and Portfolio Choice
Recent evidence suggests that lifetime experiences play an important role in determining households' investment choices. I incorporate these findings and the fact that household portfolios are underdiversified into an otherwise standard life-cycle model and examine to what extent they can help resolve long-standing puzzles in the literature regarding stock market participation and the fraction of financial wealth invested in risky assets. I show that experience-based learning about returns creates a positive correlation between a household's position in the wealth distribution and its optimism about future returns. The wealthy consequently increase their investment in risky assets, while participation is limited among poor households. I find that in a reasonably calibrated quantitative model, this mechanism is able to close approximately half of the gap between the participation rates observed in the data and the predictions from standard models. On the other hand, the average conditional risky share mostly remains unaffected.

Health Dynamics and Heterogeneous Life Expectancies *(with Jonna Olsson)*
In this paper, we provide improved estimates for age-dependent health transitions and survival probabilities for different subsamples of the U.S. population. The estimated yearly transition matrices can be used in any life-cycle model where health and survival dynamics are of interest. The results show substantial heterogeneity in life expectancy in the population. For a 70-year-old man in excellent health, the probability of reaching his 80th birthday is around 75%, while the corresponding probability for a man in poor health is just below 40%. There is also substantial inequality in life expectancy between different educational groups. In the group with less than a high school degree, the life expectancy at the age of 50 is 75 years, while the average for those with some college education or more is 80 years. This difference is due to two factors. First, at the age of 50, overall health is worse in the group with lower education. Second, even conditional on health status, the health dynamics and survival probabilities for this group are worse also from the age of 50 and onwards. We estimate that the difference in life expectancy across education groups mainly stems from the worse health and survival dynamics after the age of 50.

Subjective Life Expectancies, Time Preference Heterogeneity and Wealth Inequality *(with Jonna Olsson)*

Time preference heterogeneity is one of the potential sources of wealth inequality, but preferences are difficult to measure and quantify. In this paper, we investigate one source of time preference heterogeneity, namely heterogeneity in life expectancy, which we document using micro data. Furthermore, we document a systematic bias in subjective survival beliefs within cohorts that exacerbates the heterogeneity found in the population: individuals with a low survival probability relative to their peers underestimate their life expectancies, while individuals with a high survival probability overestimate theirs. To gauge the effect of heterogeneity in life expectancy (objective and subjective) on savings rates and ultimately wealth inequality, we introduce survival heterogeneity into an otherwise standard overlapping-generations model and let survival probabilities and beliefs evolve stochastically according to a health and death process estimated from micro data. Agents cannot insure themselves against neither health and death risks, nor against the income risk they are facing. We find strong effects of life expectancy heterogeneity on within-cohort wealth inequality, but small effects on economy-wide wealth inequality.

On the Redistributive Effects of Government Bailouts in the Mortgage Market *(with Dirk Krueger and Kurt Mitman)*

In this paper, we study the aggregate and distributional consequences of government bailout guarantees in the U.S. mortgage market. We construct a model with aggregate risk in which competitive financial intermediaries issue mortgages to households that can default on their debt. Default probabilities are priced into mortgage interest rates unless a government bailout guarantee makes the lenders whole even in economic downturns when foreclosure rates surge. We use the model to assess the extent to which bailout guarantees lead to excessive mortgage lending, household leverage, and foreclosures in episodes of housing crises, as well as excess volatility in house prices in severe recessions. We also study the moral-hazard induced misallocation between housing capital and physical capital caused by the bailout guarantee. The model features (large) aggregate shocks, housing, and capital as distinct assets and mortgage default and bailouts that are not entirely unanticipated. While we find significant effects of bailout guarantees at the individual level, the magnitude of bailouts in the aggregate is currently too small compared to historical episodes. We thus view these findings only as a first step towards assessing the welfare implications of bailout policies.

Better fast and wrong than slow and wrong.
– Anonymous Fortran programmer

Acknowledgments

The first time I heard about the IIES was in the second year of my masters, when it was time to apply for PhD programs: the faculty at the Institute for Advanced Studies (IHS) in Vienna encouraged me to send my application to Stockholm. A week later I reported that I had indeed applied to a school in Stockholm that offered programs in economics, but it turned out not to be the one they had in mind. They were thinking of the IIES, but were unaware of the fact that one could not directly apply there.

The first time I heard about Per Krusell was during a course on macro labor during my masters, where he appeared as the the central figure in "HKV," the initials of the authors whose papers on the topic we covered in the course. A few months later he came to Vienna in person, giving a talk in our seminar series. I don't remember the topic, but I recollect that I was surprised that he showed up in a t-shirt (but maybe this is just my memory playing tricks on me, projecting what I know today into the past). In any case, I do remember that I dozed off – not because the talk was boring, but because the masters program was very demanding, and afternoon seminars often turned into a semi-involuntary opportunity to get some rest. Later, during the last course I took in Vienna, we replicated a paper that was using the Krusell-Smith algorithm to solve a heterogenous-agent model. This turned out to be cool stuff, combining my interest for studying topics like inequality with technically challenging methods that involved tinkering with code, something I had enjoyed doing before turning to economics. Smith was far away, somewhere in the US, but as it happened Krusell was at the very university I had accepted to join as a PhD student. So off I went, planning to get into the IIES as soon as possible and have Per as my advisor, and I am happy that it worked out.

The first time I interacted with Kurt, who eventually became my second advisor, was during the *Quantitative Methods in Macroeconomics* course in my third year. Unsurprisingly, given my research interest, this was probably my favorite course in the program. As a young macroeconomist, it is hard not to have Kurt as a role model, given the amazing amount of well-published papers he has written in just a few years. However, I did not only benefit from his teaching and advice, I was also lucky to have him as a coauthor on a project that became the fourth chapter in this thesis.

Without Per and Kurt, this thesis would be quite different – as most PhD students, I learned a lot from interacting with my advisors. This includes small things, such as not to discard partial equilibrium modeling outright, or going back to a two-period, two-agent setting to understand an underlying mechanism even if the full model

is a quantitative mess; but also the big-picture questions which were often raised in seminars, by the entire IIES faculty. I am grateful to have had the opportunity to be part of this research environment. I want to thank participants in the macro group, and in particular Alex, Kieran, Kathrin and Tobi who provided many helpful comments at various points during my time at the institute, and especially on my job-market paper, as well as Arash, Ingvild, Jon and Robert. I also want to thank Mitch and Tessa, who helped me in their role as placement officers during my job-market year. A special shout-out goes to Roine at the Department of Economics, who at many points gave valuable feedback from a finance perspective.

If the IIES was left to researchers alone, things would quickly start falling apart, and no one would remember to celebrate the World Tapir Day. I am therefore happy that Christina was around during my whole time at the institute, and did much more than merely report on recent events in the animal sphere. I also want to thank Ulrika, especially for helping with my dissertation, as well as all those who have been part of the administrative staff over the years.

Even more than the staff, it was my fellow grad students who had a substantial impact on my daily life at the institute. Proceeding in reverse chronological order, I want to start mentioning those who ~~suffered through~~ enjoyed the job-market together with me: Has, Markus, Karin, Kasper and especially Benni, who always kept me up-to-date on any interview trick question she could find, and regularly reported on the latest news from Econ Twitter. I very much appreciate that you were around in my job-market year!

Then there are those I shared office with over the years: Sebastian, Benni, Saman (holding the record of sharing office for four years), Sreyashi, Michela, Jaakko, Hanna (a constant source of sunshine), Daniel and Sirus, as well as Mathias, who deserves special thanks for many reason, in particular for creating frequent occasions to binge on chocolate croissants and giving me the opportunity to look over his code.

Coming to the beginning of my PhD studies, I want to thank people from my cohort beyond those already listed: Anna, Eleonora, Joakim, Karl, Nadiia, Tamara, Wei, and in particular Marta.

Additionally, I want to mention those who do not fit any of the above groups, but whose presence over the last few years I appreciated: Erik, Divya (who has an impressive talent to deal with cynical people), Hannes, John, Josef, Magnus (who diligently calls me out for coming in late), Matilda, Matti, NJ, Selene, Serena, Xueping and last but not least Fabian and Philipp, wo became frequent lunch and fika companions in my final year. Thanks Fabian for newer losing sight of the big-picture questions!

Life around Stockholm University would have been very different without the 4H farm, which eventually became the destination of many after-lunch walks. I am grateful to Rosalinda and Wilmer for making these mini-excursions a source of joy.

Of course there are a few people outside of Stockholm who helped make this

possible in one way or another. I am grateful to Stefan, with whom I came up with the idea to study economics during a vacation in Italy an eternity ago. My parents Jana and Jozef supported me throughout my studies, and my sister Simona helps balance the Foltyn family by being a counterweight to an office-dwelling economist in her daily job as a journalist in the field, based in places I'd rather stay away from. I hope they will eventually come to terms with the fact that even after so many years of studying economics I am still unwilling and unable to give investment advice.

I came to Stockholm to do a PhD. I did not anticipate that it came with the benefit of meeting you, Jonna, in my second year. Over the years we have become an amazing team, which might be due to the fact that we complement each other in many areas, but also share many common interests. This is not only reflected in the two chapters in this thesis, but also in the fact that even after seven weeks of Covid-19-induced self-isolation in a single-room apartment, we are still not at each other's throats. I am glad to be moving to the Scotland with you.

Richard Foltyn
May 2020
Stockholm

Contents

Introduction

This thesis consists of four self-contained chapters, which are connected by two underlying themes. In terms of research questions, these chapters are concerned with the savings and portfolio-choice decisions households make, and the underlying drivers of such decisions. Methodologically, they share a common foundation of heterogeneous-agent modeling which has become the standard in many areas of macroeconomics and household finance. The aim of these methods is to be able to study differences in economic behavior that arise due to heterogeneity in wealth, life expectancy, beliefs, and home-ownership status, among others, as well as the differential impact of policies on such households. This connects naturally to the increasing use of micro-data to guide macroeconomic models.

In chapter 1, titled *"Experience-based Learning, Stock Market Participation and Portfolio Choice,"* I explore one possible explanation for the stylized patterns of household portfolio compositions observed in the data. These show that even in advanced economies, stock-market participation is still limited and well below 50% for most countries. Furthermore, participation is increasing in wealth, while conditional on participation, the fraction invested in risky assets is more or less flat, contrary to what most models in the household-finance literature predict.

Recent evidence suggests that lifetime experiences play an important role in determining households' investment choices. I incorporate these findings and the fact that household portfolios are underdiversified into an otherwise standard life-cycle model and examine to what extent they can help resolve long-standing puzzles in the literature regarding stock market participation and the fraction of financial wealth invested in risky assets. I show that experience-based learning about returns creates a positive correlation between a household's position in the wealth distribution and its optimism about future returns. The wealthy consequently increase their investment in risky assets, while participation is limited among poor households. I find that in a reasonably calibrated quantitative model, this mechanism is able to close approximately half of the gap between the participation rates observed in the data and the predictions from standard models.

In chapters 2 and 3, in joint work with Jonna Olsson, we investigate the heterogeneity in health and life expectancy in the U.S. population and its economic implications. Health inequality in itself is important, but to understand the underlying causes is far beyond the scope of a dissertation in economics. We therefore focus on the potential economic consequences of inequality in health and longevity. Numerous studies have

identified health dynamics and health shocks as a major source of risk over the life cycle. A negative health shock can result in large medical expenditures, which affects the incentives to accumulate assets, and could also affect the earnings potential. The survival probability directly affects the effective discount factor, a mechanism present in any life-cycle model with an uncertain life span. Some also argue that the health state directly influences the marginal utility from consumption. Hence, in order to quantify the risk an individual faces and model the choices and actions the individual takes, a realistic health and survival process is crucial. We document the health process in a structured way that is possible to use in macroeconomic models, and take a small step in thinking about how inequality in life expectancy affects economic outcomes, in particular savings.

In the second chapter, *"Health Dynamics and Heterogeneous Life Expectancies,"* we provide improved estimates for age-dependent health transitions and survival probabilities for different subsamples of the U.S. population. The estimated yearly transition matrices can be used in any life-cycle model where health and survival dynamics is of interest. The results show substantial heterogeneity in life expectancy in the population. For a 70-year-old man in excellent health, the probability of reaching his 80th birthday is around 75%, while the corresponding probability for a man in poor health is just below 40%. There is also substantial inequality in life expectancy between different educational groups. In the group with less than a high school degree, the life expectancy at the age of 50 is 75 years, while the average for those with some college education or more is 80 years. This difference is due to two factors. First, at the age of 50, overall health is worse in the group with lower education. Second, even conditional on health status, the health dynamics and survival probabilities for this group are worse also from the age of 50 and onwards. We estimate that the difference in life expectancy across education groups mainly stems from the worse health and survival dynamics after the age of 50.

In the third chapter, *"Subjective Life Expectancies, Time Preference Heterogeneity and Wealth Inequality,"* we use the results from the previous paper and ask a natural follow-up question: what are the implications of the heterogeneity in life expectancy on savings rates and ultimately wealth inequality?

According to standard economic theory a healthy person should save more for the future, all else equal, given the higher probability of living a longer life. However, an individual's consumption/savings decision is not necessarily guided by the objective statistical life expectancy, but rather by the individual's beliefs about survival. We document new facts about systematic bias in these beliefs: individuals with low survival probability relative to their peers underestimate their life expectancies, while individuals with high survival probability overestimate theirs. This systematic bias exacerbates the survival expectancy heterogeneity in the population.

To gauge the effect of survival heterogeneity, objective as well as subjective, on

inequality, we use an overlapping-generations general-equilibrium model with uninsurable idiosyncratic shocks. Agents face heterogeneous survival risk that depends on their current health state, and are subject to health shocks that follow a process estimated from data. Besides this uncertainty, we also include standard persistent and transitory shocks to labor productivity. Since we are interested in savings behavior in late life, it is important to capture other incomes during this period. Therefore we carefully model retirement benefits, closely mimicking the U.S. social security system.

We show that the standard life-cycle model gives rise to counter-factual implications when introducing survival heterogeneity. In an environment without bequests, agents with longer life expectancy save more, as expected. This is in line with the data, where individuals in better health have higher asset holdings. However, as is well known, this standard model without a bequest motive gives rise to counterfactually low savings among the elderly since agents in the model draw down their assets to virtually zero late in life. In the data, on the other hand, individuals on average have substantial asset holdings even beyond the age of 80. We therefore add a warm-glow bequest motive of the type most commonly used in the macroeconomic literature.

The effect of introducing survival heterogeneity into this environment is counter-intuitive and perhaps also unexpected: agents in poor health now save more than their healthy counterparts. The reason is as follows: since agents in poor health are more likely to die soon, they put an increased weight on bequest utility, thus raising their incentive to save. Hence, there are two effects from lower life expectancy that work in opposite directions: a shorter expected life span makes the agent save less for own consumption, but a stronger bequest motive makes the agent save more. The net effect varies depending on calibration of bequest parameters, but the second mechanism is always present with a bequest formulation of this type: a shorter life span makes agents want to save more to leave bequests. This creates a health-wealth gradient that is counter-factual and we argue that this mechanism is implausible.

We conclude that none of the standard models are adequate for investigating the effect of survival heterogeneity on savings rates and wealth inequality. We discuss possible extensions and reformulations and point out directions for further research.

Finally, in the fourth chapter, *"On the Redistributive Effects of Government Bailouts in the Mortgage Market"*, written jointly with Dirk Krueger and Kurt Mitman, we study the determinants of households' portfolio choices when these are expanded to include housing and mortgages in addition to financial assets. Furthermore, at the core of the chapter, we are interested in how government policies influence these choices to the extent that policies affect prices in the mortgage and housing markets, and thus alter households' allocation of resources between real estate and financial assets.

More specifically, we study the aggregate and distributional consequences of gov-

ernment bailout guarantees in the U.S. mortgage market. We construct a model with aggregate risk in which competitive financial intermediaries issue mortgages to households that can default on their debt. Default probabilities are priced into mortgage interest rates unless a government bailout guarantee makes the lenders whole, i.e., compensates them for losses on their mortgage portfolios in economic downturns when foreclosure rates surge. We use the model to assess the extent to which bailout guarantees lead to excessive mortgage lending, household leverage, and foreclosures in episodes of housing crises, as well as excess volatility in house prices in severe recessions. We also study the moral-hazard induced misallocation between housing capital and capital used in production caused by the bailout guarantee. The model features (large) aggregate shocks, housing, and capital as distinct assets and mortgage default and bailouts that are not entirely unanticipated.

While we do find non-negligible distributional and price effects of government guarantees at the micro level, in the aggregate the magnitude of government bailouts is too small compared to what has been observed during the Great Recession. We view these findings as merely a very first pass at a full analysis of the policy questions we are interested in. The key achievement in this chapter is thus the construction of a setting that is useful in this endeavor; our future work will hopefully demonstrate the full value of this achievement.

Chapter 1

Experience-based Learning, Stock Market Participation and Portfolio Choice

1.1 Introduction

How do households split their financial wealth among different types of assets? From U.S. data on portfolio allocations, it has long been known that wealthier households are more likely to participate in the stock market (Poterba and Samwick (1995), Haliassos and Bertaut (1995)), while the fraction invested in risky assets conditional on participation is almost flat across the wealth distribution. A similarly clear pattern emerges when examining portfolios over the life-cycle: participation is humped-shaped in age, with younger households being the least likely to invest in stocks despite their longer investment horizon. Recent evidence from Swedish and Norwegian administrative data (Bach, Calvet, and Sodini (2018), Fagereng et al. (2019)) confirms these patterns to be present in other countries and additionally documents vast return heterogeneity across the wealth distribution, which to a large extent stems from differences in exposure to risky assets.[1]

Standard models of portfolio choice, on the other hand, have largely been unsuccessful in replicating these empirical observations, in particular the limited participation (Campbell (2006), Guiso and Sodini (2013)). Early life-cycle models such as Cocco, Gomes, and Maenhout (2005) and Gomes and Michaelides (2005) more or less obtain the results established in Merton (1971), albeit in a more quantitative framework. These imply that young households, whose wealth primarily consists of non-tradable human capital (future earnings), choose to invest most if not all of their freely disposable financial wealth in risky stocks, while older households diversify their investment towards risk-free bonds. Since young households are, on average, poor in terms of financial wealth, this mechanism creates a risky share that is decreasing in wealth in the cross-section, contrary to what we observe in the data.

1. Hubmer, Krusell, and Smith (2019) and Benhabib, Bisin, and Luo (2019), in turn, establish the central role of differences in portfolio returns as a major determinant of wealth inequality in quantitative models.

More recent attempts to reconcile household-finance models with empirical findings mostly work along two dimensions: the first approach abandons the almost ubiquitous assumption of a constant relative risk aversion and instead incorporates non-homothetic utility, as in Wachter and Yogo (2010).[2] A second strand of literature imposes particular assumptions on stochastic labor income to tilt the portfolio allocation of young or poor households away from the risky asset; papers in this group include Catherine (2019) and Chang, Hong, and Karabarbounis (2018). However, their mechanisms have no effect on retired households and do not alter the extensive-margin participation decision.

In this paper, I explore an alternative explanation for the portfolio composition patterns observed in the data: based on evidence in Malmendier and Nagel (2011, 2016), the model presented here departs from rational expectations and instead lets households form beliefs about stock returns based on their history of previous realizations. Furthermore, in line with the data, I assume that households hold underdiversified portfolios, which results in idiosyncratic return histories. These two additions to an otherwise standard household-finance model give rise to positive sorting across beliefs and a household's position in the wealth distribution: rich households, to the extent that they are rich because of high returns in the past, are more optimistic and choose to invest a higher share in stocks than if they had known the true data-generating process. Poorer households who experienced low returns believe that risky returns will continue to be low, and hence shift their portfolios towards risk-free bonds, or exit the stock market altogether.

The findings from Malmendier and Nagel (2011, 2016) have recently been incorporated into models to investigate their implications for asset prices, trading volumes and return predictability (see Schraeder (2015), Nakov and Nuño (2015), Collin-Dufresne, Johannes, and Lochstoer (2016a, 2016b), Ehling, Graniero, and Heyerdahl-Larsen (2017), Malmendier, Pouzo, and Vanasco (2018), Nagel and Xu (2019)), and persistent investment slumps following severe recessions (Kozlowski, Veldkamp, and Venkateswaran 2015). These papers mainly either feature representative agents or representative cohorts in infinite-horizon settings and thus, they cannot speak to how the portfolio composition varies across the wealth distribution and over the life-cycle.

Conversely, I embed subjective belief heterogeneity into an incomplete-markets life-cycle model in which households face idiosyncratic earnings shocks. Additionally, returns on the risky asset are allowed to be imperfectly correlated within cohorts. Households know the true variance of risky returns but form beliefs about the mean excess return which they update following a rule similar to the one estimated in

2. Gomes and Michaelides (2003) introduce additive habits into a life-cycle model which can also generate a decreasing relative risk-aversion, but they find no improvement over a standard CRRA framework.

Malmendier and Nagel (2011, 2016).

I find that this mechanism can claim a partial success in reconciling portfolio choices with those observed in the data: compared to the standard model, the positive correlation between beliefs and a household's position in the wealth distribution induces limited participation even among middle-class households, something that cannot be achieved with participation costs of a reasonable magnitude. Subjective beliefs are thus able to close approximately half the gap between the participation rates observed in U.S. data and the predictions generated by a standard model. On the other hand, conditional on participation, the risky share in the cross-section mostly remains unchanged. While an individual household's optimal risky share responds to changes in beliefs in an intuitive way, these effects average out in the aggregate.

Unlike Chang, Hong, and Karabarbounis (2018) and Catherine (2019), which rely on certain properties of the stochastic earnings process to make the risky asset less desirable, the mechanism employed here continues to work for households once they retire. Furthermore, the model generates limited participation even when no participation costs are imposed. This is in contrast to the above papers and most other models where households face a positive expected excess return; in such a setting, absent any participation costs, agents choose to invest a (potentially small) positive fraction of their savings in the risky asset, irrespective of their risk aversion.

As a robustness test, I discuss an alternative model specification in which households update beliefs using Bayes' rule instead of the mechanism suggested in Malmendier and Nagel (2011, 2016). In this scenario, agents are perfectly rational (since Bayesian updating is the optimal strategy for updating beliefs as new information arrives), but suffer from limited information since they do not incorporate return realizations prior to their birth when forming beliefs. I show that the results are almost identical, even though in the Bayesian case, household expectations converge to the true excess return more quickly, thus muting the effect of subjective beliefs.

The remainder of the paper is organized as follows: In section 1.2, I discuss related papers in the literature. In section 1.3, I review the stylized facts on households' portfolio composition observed in U.S. data. Then, section 1.4 introduces a simple three-period model to illustrate the mechanism, while section 1.5 expands it to a quantitative life-cycle framework. I discuss the calibration in section 1.6 and present results for the benchmark model in section 1.7. Section 1.8 shows that these largely remain unchanged if agents use Bayes' rule to update beliefs, and section 1.9 demonstrates that the findings are robust to assuming more realistic levels of underdiversification. Section 1.10 concludes the paper.

1.2 Related literature

This paper relates to several strands of literature: first, to papers in household finance trying to explain limited stock market participation or the fraction of financial wealth invested in risky assets; second, to papers studying investors' financial decisions without imposing rational expectations; and third, to papers quantifying the importance of portfolio choice for wealth inequality. Additionally, it builds on a vast and rapidly growing empirical literature on beliefs and their implications for portfolio choice.

Household finance. Among the early seminal papers in the first group are Cocco, Gomes, and Maenhout (2005) and Gomes and Michaelides (2005).[3] As the former do not have any participation costs, in their paper household portfolio allocations do more or less reflect the results in Merton (1971): young (or poor) households invest all their financial wealth into stocks and rebalance their portfolio towards safe bonds as they age (or become more wealthy). Participation is universal for all households that choose to save a positive amount. Gomes and Michaelides (2005) extend this framework, incorporating preference heterogeneity and a fixed cost of entering the stock market. In their model, more risk-averse agents with higher precautionary savings are more likely hold stocks as they are those willing to pay the participation cost. Other papers introduce additional mechanisms to better match participation over the life-cycle: for example, Fagereng, Gottlieb, and Guiso (2017) include stock-market disaster risk, which combined with a per-period participation cost prompts older households to liquidate their stock holdings at more realistic rates.

A second group of papers proposes mechanisms that attempt to reconcile the model-generated conditional risky share with data. Catherine (2019) incorporates the earning process introduced in Guvenen, Ozkan, and Song (2014) and Guvenen et al. (2016) that exhibits cyclical skewness, i.e., the likelihood of large drops in earnings when risky returns are low, to match the portfolio composition over the life-cycle.[4] Households who are thus particularly exposed to this kind of earnings risk are more reluctant to hold a large fraction in risky assets.[5] On the other hand, Chang, Hong, and Karabarbounis (2018) use age-dependent unemployment risk with zero replacement rates and uncertainty about future earnings growth to achieve the same goal. However, once retired, households in these models revert to the same

3. In this discussion, I exclusively focus on papers with two liquid assets, a risk-free bond and a risky stock. The role of housing in households' portfolio decisions is investigated by, among others, Cocco (2004), Yao and Zhang (2004), and Vestman (2018) in a Swedish context. See Campanale, Fugazza, and Gomes (2015) for a model where the risky asset is illiquid.

4. This is an application of an idea going back to Mankiw (1986), used in Krusell and Smith (1997) and Storesletten, Telmer, and Yaron (2007), among others, to generate higher risk premia.

5. One implication is that earnings-rich households have the lowest risky share as the earnings risk enters multiplicatively in that model.

counterfactual investment choices as in the standard model, and neither mechanism affects participation.

Asset pricing and learning from experience. Another set of related papers investigates the implications of belief formation estimated in Malmendier and Nagel (2011, 2016) in an asset-pricing context. Schraeder (2015) and Malmendier, Pouzo, and Vanasco (2018) build representative-cohort models featuring CARA preferences to characterize price dynamics and household choices in the presence of non-Bayesian learning in an analytically tractable way. Collin-Dufresne, Johannes, and Lochstoer (2016b, 2016a) use a Bayesian-learning framework in which a representative agent (or two representative dynasties) are uncertain about the mean of stochastic consumption growth. In Collin-Dufresne, Johannes, and Lochstoer (2016a), the variance of the prior is reset whenever a new generation enters, thus mimicking the "learning from experience" mechanism identified in Malmendier and Nagel (2011, 2016). Collin-Dufresne, Johannes, and Lochstoer (2016a) find that this form of belief updating has a substantial effect on the the risk premium in their model. Ehling, Graniero, and Heyerdahl-Larsen (2017) introduce "learning from experience" into a Blanchard (1985)-type perpetual-youth model in which young households react more strongly to positive return shocks, thus decreasing the risk premium. Finally, in a business-cycle context, Kozlowski, Veldkamp, and Venkateswaran (2015) introduce belief updating into a representative-agent economy to generate long-lasting effects of severe recessions. In their paper, agents form beliefs about capital returns in a non-parametric fashion; new (extreme) observations can thus skew the estimated kernel density which agents use to forecast future returns for a prolonged period of time.

Portfolio choice and wealth inequality. A strand of literature that is not directly related to portfolio choice, but establishes its importance for wealth inequality, includes papers such as Hubmer, Krusell, and Smith (2019) and Benhabib, Bisin, and Luo (2019). Neither of these contains an explicit portfolio choice, but Hubmer, Krusell, and Smith (2019) incorporate exogenous heterogeneous returns into their model and find that these are a major driver of wealth inequality, in particular affecting the right tail of the wealth distribution. They calibrate a wealth-level-dependent return process to moments reported in Bach, Calvet, and Sodini (2018). However, it is important to stress that the increase in portfolio returns along the wealth gradient is primarily a result of household choices, and return heterogeneity conditional on asset class is only of secondary importance. To illustrate this, Bach, Calvet, and Sodini (2018) report that the lowest decile of the wealth distribution earns an annual excess return of 0.61% on their total financial wealth, whereas the corresponding value is above 4% for the wealthiest households. However, the return difference on *risky* financial assets is substantially smaller, since the corresponding excess returns are 5.9% for the

lowest decile and around 7.7% for the richest households. Thus, the share of risky financial assets, which rises from 10% for the poorest to almost 60%, is the main driver of return heterogeneity, and highlights the importance of understanding the determinants of household portfolio choices. Kuhn, Schularick, and Steins (2019) provide empirical evidence for the importance of portfolio composition for aggregate wealth dynamics: Using decades of data from the SCF, they show that increases in house prices reduce wealth inequality as this mostly benefits the middle class who hold leveraged positions in housing. On the other hand, gains in the stock market have the opposite effect, i.e., increasing the wealth share held by the richest since middle-class households participate in stock markets to a lesser extent.

Empirical literature on beliefs and portfolio choice. There is a vast empirical literature on beliefs about asset returns and how these relate to household portfolios. One general conclusion is that beliefs in the population are very heterogeneous and belief updating can be classified into a few distinct types that show either extrapolative or mean-reverting behavior, or evolve as if returns were a random walk (Gaudecker and Wogrolly (2019), Heiss et al. (2019), Dominitz and Manski (2011)).[6] At an aggregate level, Greenwood and Shleifer (2014) examine the average expected returns from six different surveys and find them to be highly correlated with past returns, interpreting this as evidence for extrapolative beliefs. Similarly, Vissing-Jorgensen (2003) reports that investors are more optimistic about the stock market after high returns on their own portfolios; moreover, she documents heterogeneity in beliefs depending on investors' years of experience. Numerous papers find a positive relationship between higher expected returns and the probability of participating in the stock market (Arrondel, Hector, and Tas (2014), Kézdi and Willis (2011), Hurd and Rohwedder (2012), Dominitz and Manski (2007)), even though the magnitudes are usually small. Similar findings have been established for the conditional risky share: while there is a statistically robust relationship, the effect of higher expected returns on risky shares is an order of magnitude lower than predictions from standard models (Giglio et al. (2019), Kézdi and Willis (2011), Ameriks et al. (2019)). Another group of papers does not directly observe beliefs, but relates return histories to individuals' investment decisions and finds evidence for reinforcement learning: Choi et al. (2009) report that higher past performance on retirement accounts is associated with increases in retirement savings, while Meyer and Pagel (2019) establish that the probability of reinvesting funds after quasi-random mutual fund closures is higher for investors who experienced gains. Kaustia and Knüpfer (2008) provide evidence that the likelihood of participating in IPOs decreases for investors who previously experienced poor performance. Malmendier and Nagel (2011) find that individuals who lived

6. Giglio et al. (2019) argue that beliefs are very persistent and most of the dispersion is due to between-person variation that cannot be explained by observables.

through times of high returns are more likely to participate in the stock market and invest a higher share in stocks. Lastly, Briggs et al. (2019) provide quasi-experimental evidence for the effect of wealth on stock market participation using data on Swedish lottery winners. They document that standard models with participation costs predict responses in the participation rate which are three to four times larger than those found in the data. On the other hand, they find that subjective beliefs elicited in a supplementary survey can account for about half of this gap. Additionally, splitting their sample into individuals who experienced high vs. low stock market returns prior to winning the lottery, they show that the participation responses are significantly larger within the first group.

1.3 Household portfolios in U.S. data

In this section, I document portfolio allocations in the Survey of Consumer Finances (SCF) using the 1998, 2001, 2004 and 2007 waves. I restrict the sample to include individuals aged 20 to 79. The analysis follows the one in Chang, Hong, and Karabarbounis (2018), but I additionally report portfolio composition across the wealth distribution, while they focus on the life-cycle. The variable definitions of aggregate asset categories (safe and risky financial assets) are identical to their definitions (see their appendix for a more detailed description of the SCF and the variables used).

I exclusively focus on how households invest their *financial wealth*, which does not include housing or actively managed businesses, but does include the value of businesses owned but not actively managed by households. For completeness, Table 1.5 and Table 1.6 in the appendix report detailed summary statistics for important components of households' balance sheets which also include asset and debt positions that are not part of financial wealth.

I will mostly be concerned with two statistics characterizing household portfolios across the wealth distribution and over the life-cycle:

1. The participation rate in risky assets, i.e., the fraction of households who hold any risky assets.

2. The conditional risky share, i.e., the amount of financial wealth invested in risky assets as a fraction of total financial wealth, conditional on a non-zero amount held in risky assets.

In the following graphs, I distinguish between the "gross share," which is defined as the risky share obtained when using *gross* safe assets, and the "net share" which controls for consumer debt. Using S and R to denote gross safe and risky financial assets, respectively, the gross risky share is defined as

$$\xi_{gross} = \frac{R}{R+S}$$

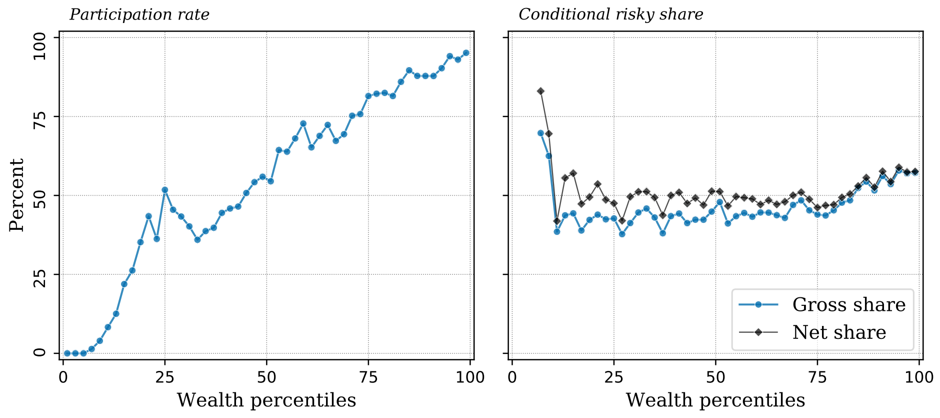


Figure 1.1: Portfolio composition along percentiles of gross total wealth. Each dot represents two percent of households. Data source: SCF 1998–2007.

while the net risky share is

$$\xi_{net} = \frac{R}{R + S - B}$$

where B is the sum of credit card and consumer loans. The difference $S - B$ is thus a measure of net safe assets.

Figure 1.1 plots the portfolio composition along gross total wealth, which is the wealth measure I use when matching model moments to data. Two stylized facts emerge: the participation rate is almost monotonically increasing from zero to close to 100% along the wealth distribution, while the conditional risky share is more-or-less flat, except for the first wealth decile. The graphs show that for the most part, the gross and net risky shares are quite close, which is due to the fact that consumer debt is not very high on average. In the appendix in section 1.A, I show that these plots are quite similar when using other wealth measures such as gross financial wealth or net worth.

In Figure 1.2, I report the evolution of the portfolio composition over the life-cycle, averaged over 5-year bins (ages 20–24, …, 75–79). The participation rate exhibits a hump-shaped pattern: only 30% of young households invest in risky assets, whereas this value peaks at around 65% for ages close to retirement. After that, participation slopes down to approximately 40% for those aged 80.

On the other hand, the conditional risky share, shown in the right-hand panel of Figure 1.2, is almost flat across age. Similar to the previous graphs, it makes little difference whether the gross or net risky share is used.

The evidence presented above ignores that fact that the most important asset on many household's balance sheet is residential real estate (the homeownership

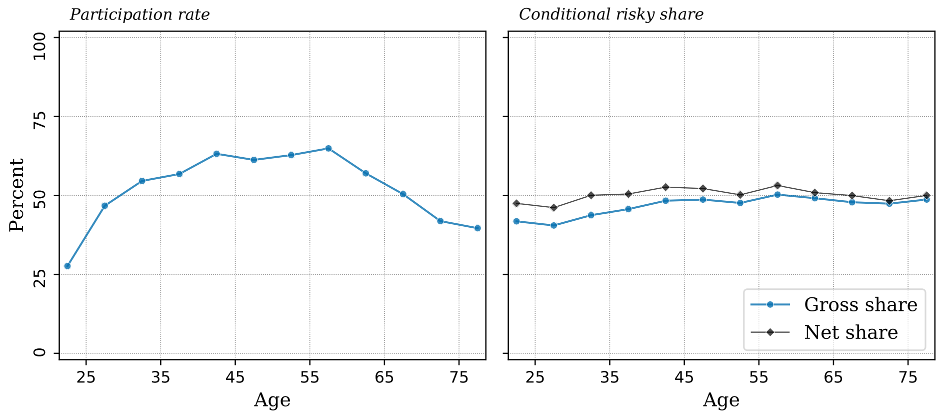


Figure 1.2: Portfolio composition over the life-cycle, 5-year averages. Data source: SCF 1998–2007.

rate is approximately 69.4% in this SCF sample). Since there is no housing in the model presented below, one potential concern is that financial wealth allocations differ systematically between home owners and renters because housing affects the investment decision of non-housing financial assets.

Figure 1.3 suggests that this is not the case: conditioning on wealth (in this case deciles of gross total wealth), homeowners and renters allocate their financial wealth similarly.[7] By construction, conditional on total gross wealth, the balance-sheet composition of homeowners and renters will look very different, so I present plots using alternative definitions of wealth in the appendix, section 1.A. The findings remain unchanged for these wealth measures.

The picture changes somewhat over the life-cycle, as shown in Figure 1.4. While the conditional risky share is almost identical for owners and renters, participation varies markedly between the two sub-populations. The reason is that renters are substantially poorer on average, which is not evident from Figure 1.3 since those moments condition on wealth. For example, in this SCF sample the median gross wealth among owners is $258,000 vs. $2,100 for renters, while the corresponding figures for net worth are $174,000 vs. $567, and $55,000 vs. $1,930 for gross financial wealth.

In Figure 1.5, I plot how the homeownership rate varies with wealth for several wealth aggregates, which illustrates a stark increase in homeownership along the wealth distribution, as expected. For example, while there are zero owners in the first decile of total gross wealth (left-hand panel), this fraction increases to almost 100% for the wealthiest. The picture is similar for various definitions of financial wealth

7. All figures plotting portfolio choices disaggregated by home ownership show the gross risky share.

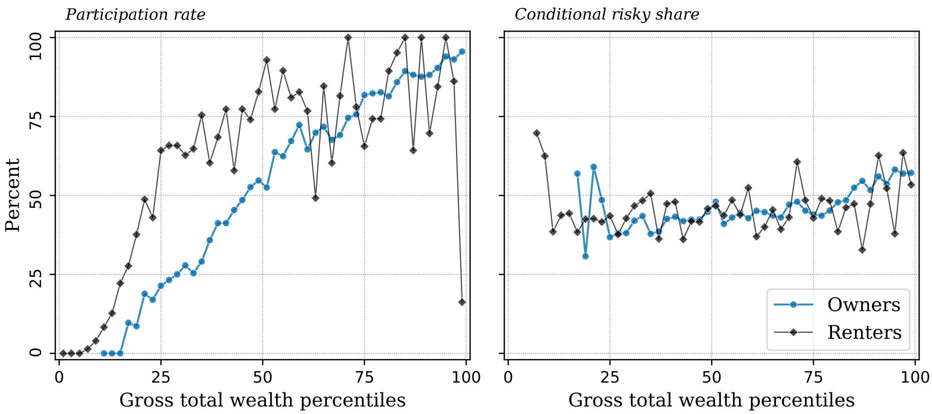


Figure 1.3: Portfolio composition along percentiles of gross total wealth by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

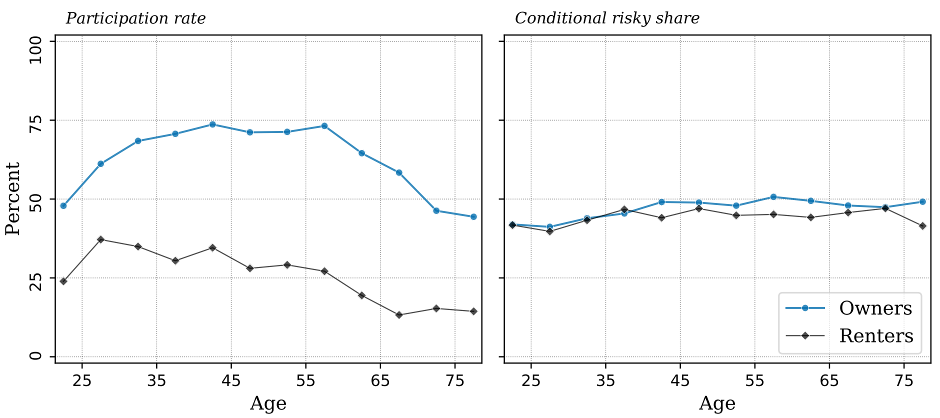


Figure 1.4: Portfolio composition over the life-cycle by homeownership status, 5-year averages. Data source: SCF 1998–2007.

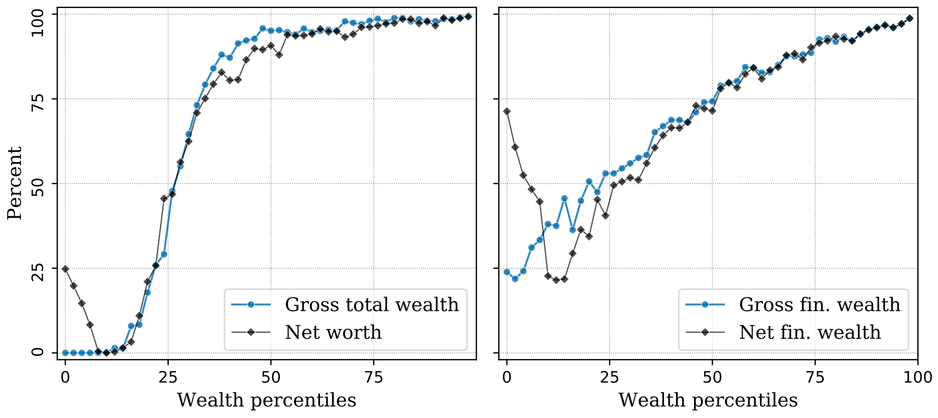


Figure 1.5: Share of homeowners along wealth percentiles. Left panel: total gross wealth and net worth. Right panel: Gross and net financial wealth. Data source: SCF 1998–2007.

(right-hand panel), where the homeownership rate is monotonically increasing in wealth except for the first decile.

To summarize the portfolio composition of homeowners vs. renters, the financial portfolios of these groups do not differ to any great extent conditional on wealth, and are particularly close when conditioning on financial wealth levels. Over the life-cycle, the conditional risky share is similar for both groups, while participation is uniformly shifted downwards for renters. This is driven by substantially lower wealth levels among renters and the fact that participation increases in wealth, as documented above.

In conclusion, the stylized facts highlighted in this section are as follows: participation substantially increases along the wealth distribution, from basically zero to 100%, while it is hump-shaped over the life-cycle, peaking around retirement. The share of financial wealth invested in risky assets, conditional on holding a positive amount of such assets, is mostly flat both along the wealth distribution and over the life-cycle.

In the remainder of the paper, I explore to what extent these patterns can be accounted for in a model with experience-based learning that gives rise to subjective beliefs about risky returns.

1.4 Simple three-period model

Before introducing the full quantitative model, I begin the exposition with a simplified three-period model which is nevertheless rich enough to illustrate the main

mechanism. The terminal period is only needed so that households face a non-trivial decision problem in the second period, as opposed to simply consuming all resources, and can otherwise be ignored.

In the first period ($t = 1$), I assume that households indexed by i are ex ante identical. They decide on consumption c, total savings b and the share invested in the risky asset, denoted by ξ. Their problem can be written as

$$\begin{aligned} V_1(a_1) &= \max_{c_1, b_1, \xi_1} \left\{ u(c_1) + \beta \mathbf{E} V_2(a_2, \widehat{\mu}_{i2}) \right\} \\ \text{s.t.} \quad a_1 &= c_1 + b_1, \qquad c_1 \geq 0, \; b_1 \geq 0 \\ a_2 &= \Big(\xi_1 R_{i2} + (1 - \xi_1) R_f \Big) b_1, \qquad \xi_1 \in [0, 1] \end{aligned}$$

where a_1 are beginning-of-period assets and $u(\bullet)$ is the standard power utility function with relative risk-aversion γ. Households can choose to invest either in a risk-free bond with gross return R_f, or in a risky asset with excess return

$$R_{it+1} - R_f = \overline{\mu} + z_{it+1}, \qquad z_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right) \tag{1.1}$$

Risky return realizations are i.i.d. in the cross-section and across time and thus, each household has a potentially different return realization R_{it+1}. Households are uncertain about the true value $\overline{\mu}$ and instead have household- and time-specific beliefs, denoted by $\widehat{\mu}_{it}$. However, the variance of risky returns σ^2 is known with certainty.

As households start out ex-ante identical with the same wealth level a_1 and the correct belief $\widehat{\mu}_{i1} = \overline{\mu}$, they choose the same portfolio allocation. Absent any labor income, the optimal choice is sufficiently well approximated by the standard formula $\xi_1 \approx \overline{\mu}/(\gamma\sigma^2)$.[8]

At the beginning of period two, after observing their risky return realizations R_{i2}, households update their beliefs according to

$$\widehat{\mu}_{i2} = (1 - \alpha)\widehat{\mu}_{i1} + \alpha \left(R_{i2} - R_f \right) \tag{1.2}$$

where α is the weight put on the most recent return realization. I postpone the discussion on the exact way in which α is determined until later; for now it suffices to say that various updating strategies (experience-based learning as in Malmendier and Nagel (2011, 2016), Bayesian learning from experience, constant-gain learning) can be mapped into a corresponding value for α.

Equation (1.2) implies that starting in period two, there will be a dispersion of beliefs in the cross-section that closely mimics the risky return distribution. The

8. The chosen preference and return parameters ensure that the optimal risky share in the first period is interior, i.e., $\overline{\mu}/(\gamma\sigma^2) \in (0, 1)$.

optimization problem in the second period thus has an additional state variable $\widehat{\mu}_{i2}$ and can be stated as

$$\begin{aligned} V_2\left(a_2, \widehat{\mu}_{i2}\right) &= \max_{c_2,\, b_2,\, \xi_2} \left\{u(c_2) + \beta \mathbf{E}_i u(c_3)\right\} \\ \text{s.t.} \qquad a_2 &= c_2 + b_2, \qquad c_2 \geq 0,\ b_2 \geq 0 \\ c_3 &= \left(\xi_2 R_{i3} + (1-\xi_2) R_f\right) b_2, \qquad \xi_2 \in [0,1] \end{aligned}$$

The problem is the same as in the first period, except that now households form subjective expectations about returns in $t = 3$, and will therefore choose different risky shares

$$\xi_{i2} \approx \begin{cases} 0 & \text{if } \widehat{\mu}_{i2} \leq 0 \\ \widehat{\mu}_{i2}/(\gamma\sigma^2) & \text{if } 0 < \widehat{\mu}_{i2} < \gamma\sigma^2 \\ 1 & \text{else} \end{cases}$$

Since there is no other heterogeneity, idiosyncratic return realizations do not only determine the beliefs in period two, but also a household's position in the wealth distribution. By construction, the wealthiest also end up being most overoptimistic about future returns, and consequently choose a higher risky share than in the fully rational model. This outcome is illustrated in Figure 1.6. By contrast, the optimal risky

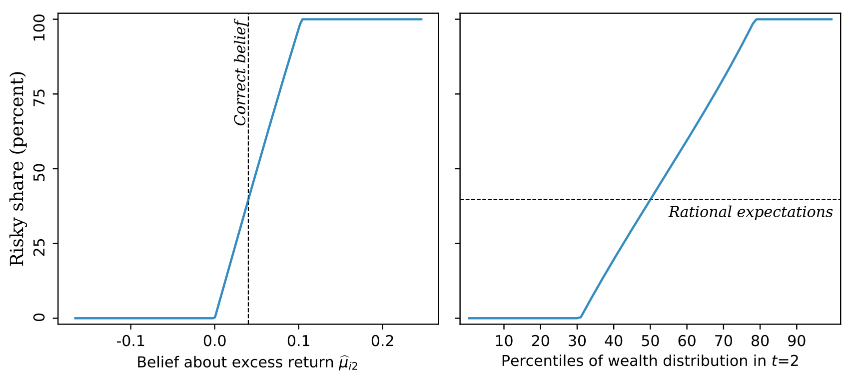


Figure 1.6: Risky shares across the belief and wealth distribution in the second period

share with rational expectations is illustrated by the dashed line, and is constant across the wealth distribution (in fact, it is the same as in the first period since the objective risky return distribution remains unchanged). Thus subjective belief heterogeneity creates an upward-sloping risky share in the cross-section and non-participation for low-wealth households. Since short-selling is not permitted, the poorest households (who at the same time are the most pessimistic about risky returns) choose not to hold any stocks, thus generating limited participation in the cross-section.

Naturally, this stylized setting vastly exaggerates the correlation between beliefs, wealth and portfolio choice. In reality, many factors other than asset returns affect a household's position in the wealth distribution, the most important being (uncertain) earnings and the life-cycle, which creates a hump-shaped age profile of asset holdings. To explore whether the mechanism outlined above carries over to a more quantitative framework which takes these aspects into account, I proceed to embed subjective beliefs into an otherwise standard life-cycle model.

1.5 Quantitative life-cycle model

Consider a discrete-time, partial-equilibrium life-cycle model along the lines of Cocco, Gomes, and Maenhout (2005), Gomes and Michaelides (2005) and many others in the household finance literature. I modify this framework along two dimensions, both of which are supported by empirical evidence: first, I impose that households form subjective beliefs about the risky asset's average return, and these beliefs are determined by each household's history of experienced risky returns. Second, I assume that households hold underdiversified portfolios with an idiosyncratic component, and they therefore differ in their return histories.

Besides their beliefs and return histories, households are heterogeneous in their wealth holdings, earnings capacity and age. They live up to a maximum of 89 years and face an age-dependent probability of death. Households supply labor inelastically, retire at an exogenously fixed age and receive a deterministic retirement income thereafter.

Markets are incomplete, so households cannot insure against earnings or survival risk. I additionally impose that households cannot borrow in either asset.

Subjective beliefs

Since subjective beliefs are the main non-standard building block of the model and work the same for both working-age and retired agents, I discuss these first.

As in the three-period example, households allocate their portfolio across a risk-free and a risky asset. Risky returns are assumed to be i.i.d. over time, and in the benchmark case I additionally impose that they are i.i.d. in the cross-section. While in the data individual returns on financial wealth are, of course, not independent of other investors' return realizations, Calvet, Campbell, and Sodini (2007) report that in Swedish register data, the idiosyncratic share of the variance of individual portfolio returns due to underdiversification is approximately 60%. Hence, there is something to be learned even from the i.i.d. assumption adopted in the benchmark case. I relax this assumption in section 1.9.

A household at age h chooses to invest in the risk-free asset with gross return R_f and a risky asset with excess return

$$r^e_{ih+1} \equiv R_{ih+1} - R_f = \overline{\mu} + z_{ih+1}\,, \qquad z_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right) \tag{1.3}$$

where the variance of risky returns σ^2 is constant and known.[9] After a return realization R_{ih} has been observed, agents update their beliefs according to

$$\widehat{\mu}_{ih} = (1 - \alpha_h)\widehat{\mu}_{ih-1} + \alpha_h \left(R_{ih} - R_f\right) \tag{1.4}$$

where $\widehat{\mu}_{ih-1}$ was the prevailing belief in the last period. The choice of learning mechanism determines how the weight α_h on the most recent observation is determined. I discuss three potential alternatives in turn.

Constant-gain learning. In this setting, $\alpha_h = \alpha$ is a time- and age-invariant parameter. This method is used by Nagel and Xu (2019) in a modified Bayesian framework in an infinite-horizon model. The findings in Malmendier and Nagel (2011, 2016), however, suggest that new information is assigned different weights depending on an individual's age, and are thus incompatible with a constant α.

Bayesian learning from experience (BLE). Bayes' rule is the optimal algorithm to update beliefs when new information arrives and hence, this approach is closer to rational expectations.[10] However, agents only use information from their own experience to update their beliefs, thus discarding events prior to their birth. With a sufficiently long sequence of returns, agents' beliefs would converge to the true average excess return $\overline{\mu}$ and any subjective belief heterogeneity would consequently disappear in the long run. However, in a life-cycle setting, it is reasonable to assume that newborns do not factor in information that predates their own lifetime and instead impose Bayesian learning from experience.

Experience-based learning (EBL). A third belief formation method, proposed and estimated in Malmendier and Nagel (2011, 2016) and applied in an asset pricing context in Malmendier, Pouzo, and Vanasco (2018), postulates that the update weight

9. For the remainder of the paper, I adopt the convention to subscript objects that only depend on age by h instead of t which is used to denote calendar time. In the benchmark model with i.i.d. returns in the cross-section and no time-varying aggregates, the household problem can be fully characterized in terms of household age h. Calendar time t will only play a role in section 1.9 when aggregate market returns are introduced into the model.

10. I adopt the terminology used in Malmendier, Pouzo, and Vanasco (2018) who label the belief updating methods discussed here as "Bayesian learning from experience" and "experience-based learning."

is a function of age only.[11] Malmendier and Nagel (2011) show that an index of past returns, computed as

$$\overline{R}_{it}(\lambda) = \sum_{k=1}^{age_{it}-1} w\left(age_{it}, k; \lambda\right) R_{t-k} \tag{1.5}$$

(using their notation), with the weighting function given by

$$w(age_{it}, k; \lambda) = \frac{(age_{it} - k)^{\lambda}}{\sum_{j=1}^{age_{it}-1}(age_{it} - j)^{\lambda}} \tag{1.6}$$

is related to households' risk attitudes, stock market participation and the share invested in stocks conditional on participation. The functional form is sufficiently flexible to allow for past observations to receive more or less weight as one moves back in time, and collapses to the conventional estimate of the sample mean for $\lambda = 0$.

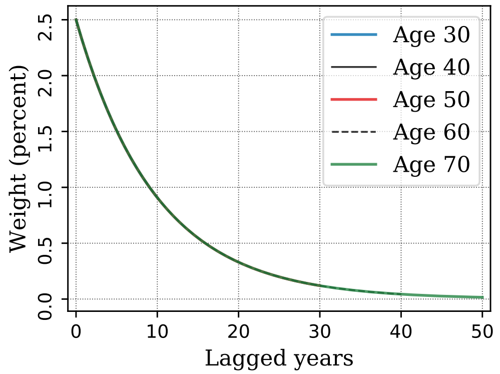
In the appendix, section 1.B, I show how the weighting scheme in (1.5) and (1.6) can be transformed into a recursive formulation and stated in terms of the belief updating equation (1.4). Adapted to the notation used in my model, the update weight is given by

$$\alpha_h = \frac{(\underline{h} + h - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h-1}(\underline{h} + h - k)^{\lambda}} \tag{1.7}$$

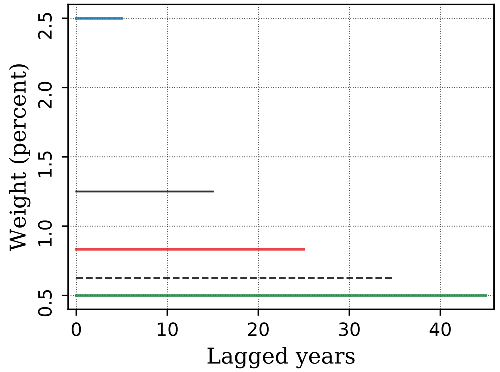
where $\underline{h}$ is the actual age corresponding to model age $h = 0$, i.e., $age = \underline{h} + h$.

Figure 1.7 shows the weights assigned to lagged observations implied by each of the three belief updating methods. The parameters are calibrated such that individuals at the age of 30 update their beliefs by approximately 2.5% (at a quarterly frequency) in all three cases. This is line with the estimates of λ in Malmendier and Nagel (2011) which at age 30 imply an update weight of about 10% at an annual frequency. The three mechanisms have starkly different implications for how past information is used (or discarded) to form beliefs: with constant-gain learning, lagged observations are weighted in a geometrically decreasing fashion, while with Bayesian learning from experience, all observations receive the same weight (since the returns are i.i.d.), even though this weight decreases as an individual collects longer time series of data. Lastly, the estimates in Malmendier and Nagel (2011, 2016) suggest that people aged 30 put a higher weight on the most recent observation (approximately 2.5%), while

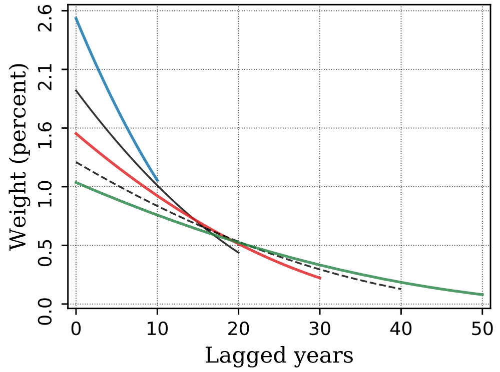
11. I combine the updating methods in Malmendier and Nagel (2011) and Malmendier and Nagel (2016) even though they use different functional forms. The latter paper assumes that people update their beliefs using a decreasing-gain learning algorithm where the gain is a function of age and birth year. However, as the authors show in the appendix to Malmendier and Nagel (2016), there is a direct mapping between the gain estimated in Malmendier and Nagel (2016) and the shape parameter λ of the weighting scheme in Malmendier and Nagel (2011), here restated in (1.5) and (1.6).



(a) Constant-gain learning



(b) Bayesian learning from experience (BLE)



(c) Experience-based learning (EBL)

Figure 1.7: Weights assigned to past observations for constant-gain learning, Bayesian learning from experience and experience-based learning.

by the age of 70 this value drops to 1%. The effect of new information is thus muted for older individuals.

In this paper, I choose experience-based learning as the benchmark case for two reasons: first, Malmendier and Nagel (2011, 2016) provide evidence that historical observations weighted in this way have a significant impact on how people form beliefs and make financial decisions. Second, EBL with $\lambda = 0$ can replicate the same weights as BLE (for some initial prior), but $\lambda = 0$ is rejected by the findings in Malmendier and Nagel (2011, 2016).

In the next section, I describe how subjective beliefs are incorporated into the household's optimization problem.

Retired households

Households exogenously retire at age H_r and continue to live up to a maximum age H, facing a stochastic survival risk along the way. Once retired, they receive retirement benefits that are proportional to their last pre-retirement labor income, which is summarized by the state p and remains unchanged through the retirement.

A retired household's state is represented by the tuple $\boldsymbol{x} = (h, a, p, \widehat{\mu}_i, j)$, where h denotes age, a is beginning-of-period cash-at-hand (i.e., any assets plus retirement benefits), $\widehat{\mu}_i$ is the current belief about excess returns and j indexes a (fixed) preference type. The household optimally chooses consumption c, total savings b and the share invested in the risky asset ξ in order to maximize

$$V_{jh}^r(a, p, \widehat{\mu}_i) = \max_{c, b, \xi} \left\{ c^{1-\psi} + \beta_j \left[\pi_h^s \mathbf{E}_i \left[\left(V_{jh+1}^r \right)^{1-\gamma} \right] + (1 - \pi_h^s) \mathbf{E}_i \left[\left(V_j^b \right)^{1-\gamma} \right] \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

with continuation values

$$V_{jh+1}^r \equiv V_{jh+1}^r \left(a', p, \widehat{\mu}_i' \right) \qquad\qquad V_j^b \equiv V_j^b \left(a_b' \right)$$

subject to the budget constraint

$$a = c + b + \kappa \cdot \mathbf{1}_{\{\xi > 0\}} \tag{1.8}$$

and the usual non-negativity constraints $c \geq 0$, $b \geq 0$. Moreover, households are not allowed to take short positions in either asset, restricting the risky share to $\xi \in [0, 1]$. I assume that the preferences are of the Epstein-Zin-Weil form (Epstein (1988) and

Epstein and Zin (1989), Weil (1990)) such that the EIS ψ^{-1} is not restricted to be the inverse of the relative risk aversion γ. Decoupling the EIS from the risk aversion is helpful when trying to match the wealth distribution (and hence savings).

To this end, households are also permitted to be heterogeneous in terms of preferences. These differences are assumed to be randomly drawn at birth and remain unchanged throughout their lifetime. A household's preference type, indexed by j, then determines its discount factor β_j and the bequest utility weight ϕ_j, which is discussed below. I assume that there are two types, with type 1 being the impatient household which also has a lower bequest motive.

In the benchmark calibration, I impose a fixed per-period participation cost κ when households choose to invest a positive amount in the risky asset, but additionally report results for $\kappa = 0$.

Given optimal choices and next-periods shock realizations, the ex-post return on a household's portfolio is

$$R_p' = \xi(R_i' - R_f) + R_f$$

and hence next-period cash-at-hand is

$$a' = R_p' b + \rho_{ss} p$$

with ρ_{ss} being the replacement rate relative to the earnings level p just prior to retirement.

Households survive with an age-dependent survival probability π_h^s with $\pi_H^s = 0$ in the terminal period. Upon death, they derive utility from "warm-glow" bequests as in De Nardi (2004), summarized by the function[12]

$$V_j^b(a_b) = \phi_j^{1/(1-\psi)} a_b \tag{1.9}$$

The parameter ϕ_j scales the utility derived from bequests relative to current-period consumption and the continuation value conditional on survival. While it is common in the macroeconomic literature to include a non-homothetic luxury-good component in the bequest motive such that it predominantly affects wealthy households, this induces a relative risk aversion that is increasing in wealth, which is undesirable in a context studying portfolio choice.

Subjective beliefs enter the household's problem via the state variable $\widehat{\mu}_i$, which I subscript by i to make explicit the link between household i's beliefs about mean excess returns and its subjective expectations, denoted by $\mathbf{E}_i[\dots]$. In the benchmark

12. This simple linear function can be obtained from a more general functional form for the case of EZW preferences suggested in Bommier, Harenberg, and Le Grand (2017) by setting all their additive terms to zero (in particular the bequest shifter). With power utility, when $\gamma = \psi$, the bequest utility implied by (1.9) is the familiar $v_j^b(a_b) = \phi_j a_b^{1-\gamma}/(1-\gamma)$.

model, I assume that households are naive with respect to any future belief updating, i.e., they do not anticipate that they will revise their beliefs about risky returns as they observe realizations in the future. Therefore, their perceived law-of-motion for $\widehat{\mu}_{ih}$ is $\widehat{\mu}_{ih+1} = \widehat{\mu}_{ih}$. Alternatively, households could be sophisticated and factor in that for a given realization of R'_i tomorrow, they will update their beliefs according to (1.4) with the weight on the last observations given by (1.7). This difference arises only in terms of how expectations are formed; when simulating households, their beliefs are always updated according to (1.4). I report results for an economy with sophisticated households in the appendix, since for the benchmark model, the differences are modest.

Working-age households

Working-age households are assumed to supply labor inelastically and therefore, their optimization problem is almost identical to the one discussed above. An agent with a state vector given by $\boldsymbol{x} = (h, a, p, \widehat{\mu}_i, j)$ for $0 \leq h < H_r$ solves

$$V_{jh}\left(a, p, \widehat{\mu}_i\right) = \max_{c,\, b,\, \xi} \left\{ c^{1-\psi} + \beta \left[\pi_h^s \mathbf{E}_i \left[\left(V_{jh+1} \right)^{1-\gamma} \right] + (1 - \pi_h^s) \mathbf{E}_i \left[\left(V_j^b \right)^{1-\gamma} \right] \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

with continuation values

$$V_{jh+1} \equiv \begin{cases} V_{jh+1}\left(a', p', \widehat{\mu}'_i\right) & \text{if } h < (H_r - 1) \\ V^r_{jh+1}\left(a', p, \widehat{\mu}'_i\right) & \text{else} \end{cases}$$

and

$$V_j^b \equiv V_j^b\left(a'_b\right)$$

The constraints are identical to those of the retired household. However, conditional on survival, next-period cash-at-hand is now given by

$$a' = R'_p b + y'$$

where y' are stochastic earnings. I adopt the standard convention of modeling these as

$$\log y_{ih+1} = \log \omega_{h+1} + \log p_{ih+1} + \log \epsilon_{ih+1} \tag{1.10}$$

where ω_h is a hump-shaped, deterministic age-profile that reflects the average earnings growth over the life-cycle. Idiosyncratic uncertainty enters via the persistent labor component p_{ih}, which follows an AR(1) in logs,

$$\log p_{ih+1} = \rho_p \log p_{ih} + \nu_{ih+1}, \qquad \nu_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(-\frac{1}{2}\sigma_\nu^2,\, \sigma_\nu^2\right)$$

as well as the purely transitory shock ϵ_{ih}, distributed as

$$\log \epsilon_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(-\frac{1}{2}\sigma_\epsilon^2,\, \sigma_\epsilon^2\right).$$

I normalize the average wage level to unity and omit it from earnings altogether.

Belief updating and subjective expectations are identical to the retired household's problem. The only additional assumption needed is the distribution of beliefs of newborns with $h = 0$, which I discuss in the calibration section.

1.6 Calibration

Externally calibrated parameters

Demographics

Households enter the economy at the age of 20 and live for a maximum of 70 years up to a maximum age of 89. They exogenously retire after the initial 45 years. The age-dependent survival probabilities are taken from the U.S. period life tables (Arias and Xu 2018) for the year 2015, using the values for the white male sub-population. For the quarterly calibration, I assume that the survival probabilities are constant within a year and given by $\pi_h^{s,qtr} = \left(\pi_h^{s,annual}\right)^{1/4}$. The implied life expectancy at birth is 80.4 years (i.e., 60.4 model years).

Labor income

For the deterministic age profile of earnings $(\omega_h)_{h=0}^H$ in (1.10), I use the estimates for the high-school-educated sub-sample reported in Cocco, Gomes, and Maenhout (2005). This profile also includes the replacement rate ρ_{ss} for retirement benefits, which Cocco, Gomes, and Maenhout (2005) estimate to be approximately 68% for this group.

I take the parameters (ρ_p, σ_ν^2) that govern the persistent earnings component as well as the variance of the transitory shocks σ_ϵ^2 from Krueger, Mitman, and Perri (2016) which they estimate from the PSID.

These parameter values are listed in Table 1.1. I normalize the (cohort-size-weighted) deterministic earnings profile for working-age households to one, such that the average earnings of working-age households are always unity.

Parameter	Description	Value	Source
Demographics			
$\underline{h}$	Initial age (in years)	20	
H	Maximum attainable age (in years)	89	
H_r	Retirement age (in years)	65	
π^s_h	Survival probabilities	–	US life tables (2015)
Earnings			
ρ_p	Auto-correlation of persistent earnings	0.9695	Krueger, Mitman, and Perri (2016)
σ_v	Std. dev. of persistent earnings shock	0.1960	Krueger, Mitman, and Perri (2016)
σ_ϵ	Std. dev. of transitory earnings shock	0.2284	Krueger, Mitman, and Perri (2016)
ω_h	Age-dependent earnings profile	–	Cocco, Gomes, and Maenhout (2005)
ρ_{ss}	Retirement income replacement rate	0.6828	Cocco, Gomes, and Maenhout (2005)

Table 1.1: Demographic and earnings parameters (annual)

Parameter	Description	Value	Source
Returns			
R_f	Gross risk-free return	1.02	Cocco, Gomes, and Maenhout (2005)
$\overline{\mu}$	Risk premium	0.04	Cocco, Gomes, and Maenhout (2005)
σ	Volatility of risky returns	0.16	Cocco, Gomes, and Maenhout (2005)
Subjective beliefs			
λ	Belief updating weight	1.5	Malmendier and Nagel (2011)
–	Cross-sectional mean of initial beliefs	0.04	–
–	Cross-sectional std. dev. of initial beliefs	0.09	–

Table 1.2: Risky return and belief formation parameters (annual)

Beliefs and returns

I adopt the standard values for the risk-free interest rate, the risk premium and the volatility of risky returns used in the household-finance literature. These are reported in Table 1.2.

Turning to beliefs, Malmendier and Nagel (2011) estimate the shape parameter λ for the experience-based learning model in (1.5), which governs the weighting of past realizations, for several outcome variables including stock market participation and the share invested in stocks conditional on participation. Their point estimates are in the range of approximately 1.3–1.9 at an annual frequency, depending on the exact specification. I choose λ to be 1.5 for an annual model, which implies that a 30-year-old individual assigns a weight of about 10% to the most recent observation. For the quarterly model, I therefore choose $\lambda = 2.0$ so that households aged 30 update their beliefs by $\approx 2.5\%$ as new information arrives.

The initial distribution of beliefs used to simulate the economy is determined as follows: I assume that households observe risky returns for five years (20 periods in the quarterly model) before becoming economically active and making investment

choices. Their initial belief distribution five years prior to entering the economy is identical to the risky return distribution (as they only have this single observation), which they then update according to the algorithm described above. This implies that when agents become economically active at the age of 20, the cross-sectional distribution of their beliefs about excess returns has a standard deviation of approximately 0.09 at an annual or 0.037 at a quarterly frequency.[13]

Since excess returns are Gaussian and the updating rule prescribes a deterministic (age-dependent) weight to be applied to new observations, this allows me to exactly characterize the distribution of beliefs at each age. As the initial distribution is a linear combination of Gaussian realizations, it is itself Gaussian at age $h = 0$. Applying the belief updating rule in (1.4), the cross-sectional distribution at any age $h + 1$ is therefore again Gaussian with a mean that evolves according to

$$\mathbf{E}\widehat{\mu}_{ih} = (1 - \alpha_h)\mathbf{E}\widehat{\mu}_{ih-1} + \alpha_h \mathbf{E}\left[R_{ih} - R_f \right] = (1 - \alpha_h)\mathbf{E}\widehat{\mu}_{ih} + \alpha_h \overline{\mu} \tag{1.11}$$

where the expectation is taken over the distribution of households aged h. The law of motion for the cross-sectional variance of within-cohort beliefs is

$$\begin{aligned} \mathrm{Var}\left(\widehat{\mu}_{ih} \right) &= (1 - \alpha_h)^2 \, \mathrm{Var}\left(\widehat{\mu}_{ih-1} \right) + \alpha_h^2 \, \mathrm{Var}\left(R_{ih} - R_f \right) \\ &= (1 - \alpha_h)^2 \, \mathrm{Var}\left(\widehat{\mu}_{ih-1} \right) + \alpha_h^2 \sigma^2 \end{aligned} \tag{1.12}$$

The evolution of beliefs in the cross-section is plotted in Figure 1.8. As equations (1.11) and (1.12) suggest, if newborns' beliefs are centered around the true excess return, beliefs will on average remain unbiased at any age. Furthermore, the within-cohort variance of beliefs collapses as a cohort ages, albeit at a slower pace than if agents had updated their beliefs optimally, which I discuss in section 1.8 in the context of Bayesian learning from experience.

Parameters determined by moment matching

I use the remaining preference parameters to approximately match the wealth distribution relative to average earnings as obtained from the SCF.[14] To this end, during the simulation newborn households draw initial wealth levels that correspond to the wealth distribution observed in the SCF for ages 20–25. This, however, turns out not to matter much in this class of models as in the presence of an upward-sloping earnings profile young households optimally consume their assets, thus leveling any differences in initial endowments.[15]

13. These differences stem from the fact that λ depends on the period length, and households have four times as many observations in the quarterly model.

14. As in Chang, Hong, and Karabarbounis (2018) who use the same SCF sample, the average earnings in the data are assumed to be \$40,000 per year.

15. If permitted to do so, young households would in fact prefer to borrow against their future income to smooth lifetime consumption.

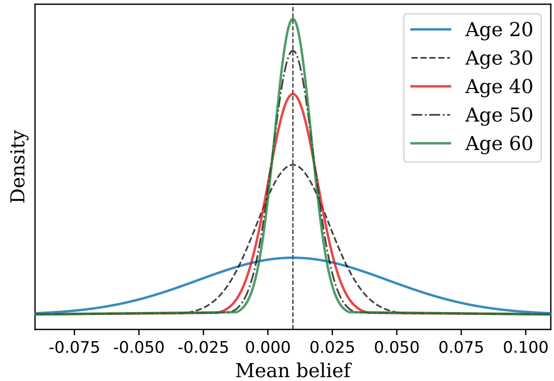


Figure 1.8: Evolution of beliefs in the cross-section with experience-based learning as in Malmendier and Nagel (2011, 2016).

Parameter	Description	Value	Source
γ	Relative risk aversion	5	Standard
β_j	Discount factor	0.88, 1.06	–
ψ^{-1}	Elasticity of intertemporal substitution	0.35	–
ϕ_j	Weight on bequest utility	0.45, 10.0	–
–	Fraction of impatient households	0.85	–

Table 1.3: Preference parameters (annual)

The relative risk aversion γ is assumed to be 5 for both household types, as a value of this magnitude is commonly used in the recent literature (Chang, Hong, and Karabarbounis (2018), Catherine (2019), Vestman (2018)).[16] The remaining parameters $\{\beta_j\}_j$, $\{\phi_j\}_j$, the elasticity of intertemporal substitution ψ^{-1} and the fraction of impatient households are set to values reported in Table 1.3.

A few comments are in order: while the discount factor heterogeneity seems large compared to calibrations in other contexts such as Krusell and Smith (1998), who were the first to introduce β-heterogeneity into a quantitative model, these magnitudes are required to generate a plausible level of wealth inequality in this setting. The reason is that in life-cycle models which already have a great deal of background risk, heterogeneity in the discount factors is less powerful (see Hendricks (2007), and Foltyn and Olsson (2019) for a setting in which discount factor heterogeneity

16. Setting a higher value such as $\gamma = 10$, which is also a commonly used value, makes matching the wealth distribution almost impossible as it induces substantial precautionary savings in the presence of highly persistent earnings risk. Thus the "poor" households end up being substantially richer than in the data.

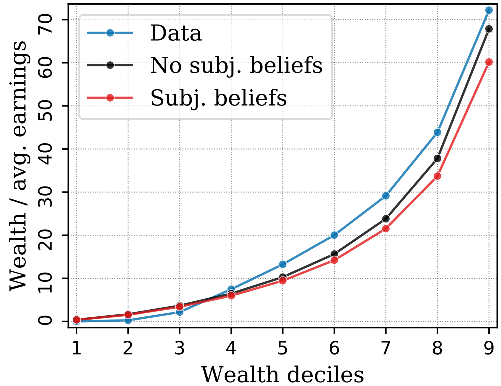


Figure 1.9: Wealth distribution in SCF vs. model by wealth decile. Each dot shows the mean wealth conditional on being in a given decile, normalized by average quarterly earnings. Data source: SCF 1998–2007.

results from differences in life expectancy). In contrast, in Krusell and Smith (1998), households only face a low-persistence unemployment risk. Furthermore, without intergenerational wealth transfers, as in the present model, any wealth built up by a cohort vanishes as that generation exits the economy, thus preventing the accumulation of substantial wealth which would be the case in infinite-horizon models such as Krusell and Smith (1998).

Turning to the heterogeneity in the weight ϕ_j that households assign to their "warm-glow" bequest utility, the differences seem large in the context of CRRA utility. However, with EZW preferences, these values interact with the EIS and γ, so the effective difference for this parametrization is much smaller.

Figure 1.9 compares the resulting wealth moments from the model to those obtained from the SCF. The match is reasonably good, even though the model struggles to generate the wealth concentration at the very top. In the tenth decile, which is not shown as it makes the differences in the lower parts of the distribution hard to read, the model can generate only about 60% of wealth holdings observed in the data.

1.7 Results

Before reporting the results for the benchmark calibration with per-period participation costs, I first discuss the model *without* participation costs to illustrate to what extent subjective beliefs can generate limited participation on their own.[17]

17. In the appendix, section 1.C additionally shows the effect of introducing subjective beliefs using an alternative calibration along the lines of Cocco, Gomes, and Maenhout (2005), which has become

No participation costs

Starting with the model without participation costs, Figure 1.10 shows the average simulated household portfolio allocations in the cross-section and Figure 1.11 plots the corresponding graphs along the life-cycle. While neither model comes close to the data moments, these figures illustrate the mechanism and the improvements due to subjective beliefs.

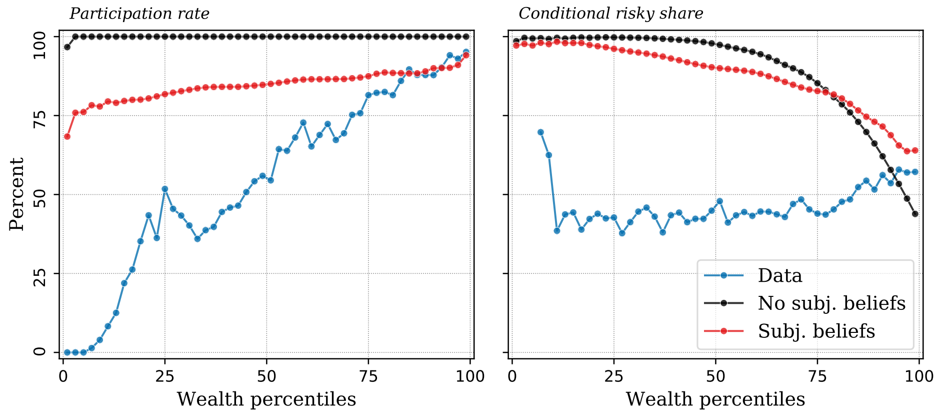


Figure 1.10: Portfolio composition along the wealth distribution. Calibration *without* participation costs.

First, subjective beliefs generate limited participation in the risky asset *without* imposing any participation costs. In contrast, as is well-known, the participation rate conditional on saving a positive amount is 100% in the “standard” model. Introducing subjective beliefs pushes this value down to about 70% in the first wealth decile, and participation is upward-sloping along the wealth distribution. The conditional risky share is slightly tilted since poor households no longer choose to invest all their savings in stocks, while the wealthy increase their stock holdings. The model with subjective beliefs performs somewhat better than the rational-expectations variant, lowering the conditional risky share for households with fewer assets, and increasing it in the right-hand tail of the distribution.

Over the life-cycle, Figure 1.11 shows that the differences between the fully-rational and the subjective-beliefs models are again predominantly on the extensive margin, as younger households are less likely to enter the stock market. Compared to the rational-expectations model where almost all households choose to hold some risky assets, this figure drops to 65% for those aged 20–25.

a benchmark paper in the household-finance literature. I argue that this parametrization generates wealth levels that are highly counterfactual and is thus not well-suited to investigate portfolio choices

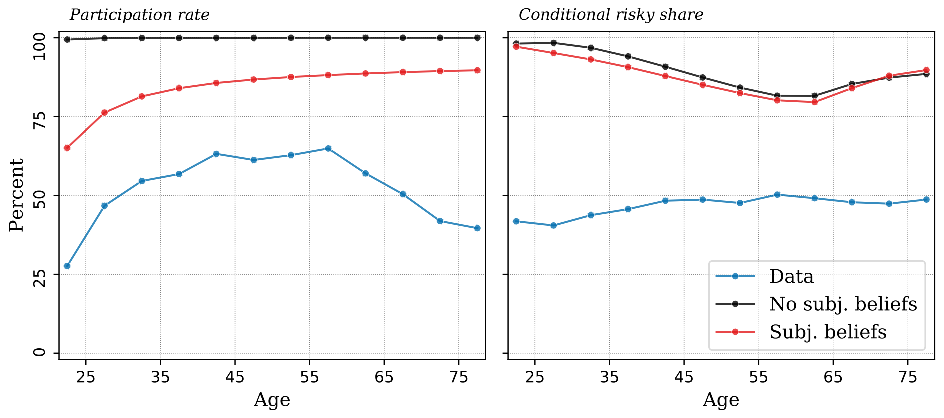


Figure 1.11: Portfolio composition over the life-cycle. Calibration *without* participation costs.

Benchmark model with participation costs

Since subjective beliefs alone do not bring participation down to the levels observed in the data, a natural follow-up question is to what extent fixed participation costs improve the model fit, and how much the effects differ from a fully-rational model in that scenario. I therefore set κ, the fixed per-period participation cost that enters the household budget constraint in (1.8), to $\kappa = 0.01$, i.e., to 1% of average per-period earnings (this amounts to around \$500 annually).

Figure 1.12 illustrates that participation along the wealth distribution in the model with subjective beliefs improves considerably, and more so than in the standard model with the same participation costs. In the latter setting, the fixed cost has no effect beyond the second wealth decile, while it pushes participation down towards its empirical counterpart in the presence of subjective beliefs. The findings for the fully-rational model are in line with previous literature on participation in the presence of fixed costs: modest fixed participation costs only affect poor households with little financial wealth, but they cannot rationalize limited participation among middle-class households. For example, Vissing-Jorgensen (2003) reports that when attempting to match the participation rates in each wealth decile using heterogeneous participation costs, she finds that the required median cost ranges from \$650–\$1,450 (in 2003 USD), depending on the PSID wave. While not being directly comparable, Briggs et al. (2019) estimate that a one-time entry cost of more than \$30,000 is needed in an otherwise standard life-cycle model to rationalize the (low) participation response among lottery winners. In comparison, in the presence of subjective beliefs, households who experienced low returns in the past expect low or even negative returns in the future,

along the wealth distribution.

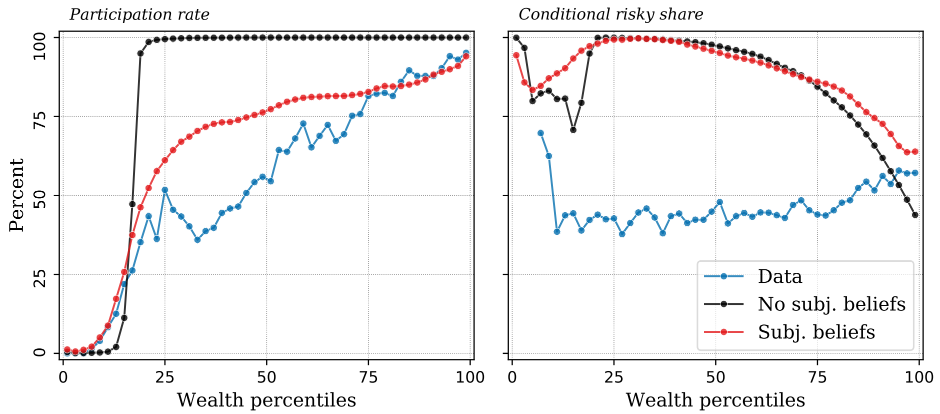


Figure 1.12: Portfolio composition along the wealth distribution. Benchmark calibration *with* participation costs.

and thus small or even no participation costs are sufficient to exclude them from the stock market.

Whereas introducing participation costs on top of subjective beliefs substantially improves the fit with data compared to a standard model, both models produce qualitatively similar predictions when it comes to the conditional risky share. Between the 20th and 70th percentiles of the wealth distribution, the risky share is almost identical for the rational-expectations and subjective-belief models, while in the latter setting it is somewhat higher for the wealthiest households since these are on average more optimistic about stock returns.

A similar picture emerges over the life-cycle, shown in Figure 1.13: both the fully-rational and the subjective-beliefs model see improvements in the participation rates along the age dimension, but the latter model is considerably closer to the data. However, the conditional risky share is again very similar in both scenarios.

Mechanism: positive sorting over beliefs and wealth

The model with subjective beliefs helps explain the limited stock market participation because it induces positive sorting across beliefs and a household's position in the wealth distribution. Consider a poorer household: one reason why a household has less wealth in this model is that it experienced repeated low returns on its investment in the past. These low return realizations at the same time induced the household to revise its beliefs about average stock returns downward, and it consequently chooses to decrease the fraction of financial wealth invested in the risky asset, or stays out of the stock market altogether. The opposite holds for a wealthy household.

In the benchmark model, earnings uncertainty and life-cycle savings behavior

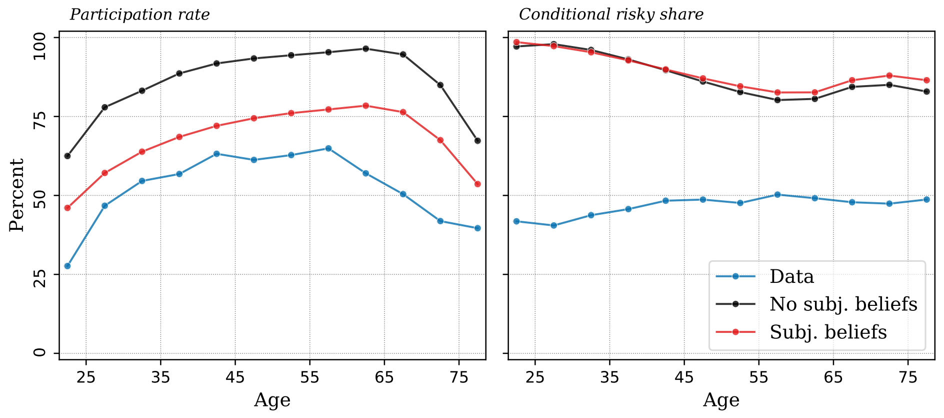


Figure 1.13: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs.

also determine a household's position in the wealth distribution in addition to past returns, unlike in the stylized example in section 1.4. There might very well be *richer* households who are *more pessimistic* about risky returns because these households are wealthy due to high earnings, or due to being close to retirement when the life-cycle asset profile peaks, i.e., for reasons unrelated to their realized stock market returns. However, *on average*, wealthier households are more optimistic about the stock market than poorer ones.

To make this point, Figure 1.14 plots the CDF over wealth for ages 20,…, 80 for selected percentiles of the subjective belief distribution.[18] At the age of 20, the sorting mechanism is not yet operational since newborns' initial wealth levels and beliefs are assumed to be uncorrelated. As households grow older, their position in the wealth distribution diverges conditional on their belief. By the age of 80, households that are in approximately the 75th percentile of the belief distribution are substantially wealthier than those less optimistic about excess returns: compared to a household with median beliefs, their CDF is uniformly shifted to the right.

While subjective beliefs due to learning from experience induce a positive correlation between beliefs and wealth in the population, for an individual household the shape of the risky share policy function is qualitatively the same as in the standard model: as shown in Figure 1.15, households diversify away from the risky asset as their cash-at-hand increases, thereby creating the usual downward-sloping optimal risky share. However, beliefs shift a household's optimal risky share up or down, with more optimistic households choosing higher risky shares, while pessimistic households

18. Given that newborn households start with subjective beliefs which are centered on the true excess return in the cross-section, the median household shown in Figure 1.14 has the correct belief at any age.

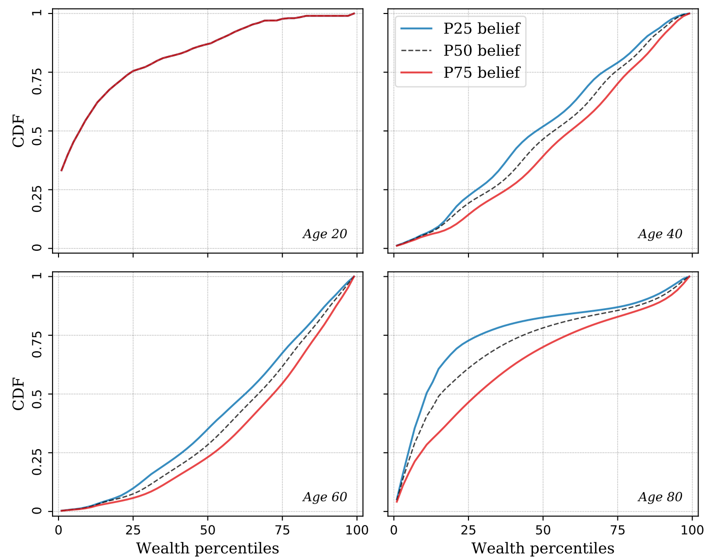


Figure 1.14: CDF over wealth for selected ages and percentiles of the subjective belief distribution. Each line shows the distribution of wealth conditional on age and beliefs. Wealth percentiles on the x-axis are computed for the overall population.

(here illustrated by the 25th percentile of the belief distribution) choose to not invest in stocks at all. Even though beliefs have large effects on individual behavior, these approximately average out in the aggregate, yielding an average conditional risky share that is not too different from the standard model.

I end this section with a more technical discussion on why the model with subjective beliefs creates limited participation even without imposing any participation costs.To simplify the exposition, I consider a one-period portfolio-choice problem with CRRA preferences and no participation costs, but the same reasoning carries over to the full model. Household i in this case maximizes

$$\max_{\xi} \; \mathbf{E}_i \left[\frac{\left(aR^p_{ih+1} + y_{h+1}\right)^{1-\gamma}}{1-\gamma} \right]$$
$$\text{s.t.} \quad R^p_{ih+1} = \xi\left(R_{ih+1} - R_f\right) + R_f \,, \qquad \xi \in [0,1]$$

where $a > 0$ is cash-at-hand, y_{h+1} is some (stochastic) realization of non-financial income tomorrow and R^p_{ih+1} is tomorrow's return on the household's portfolio. Denoting the Lagrange multipliers on the constraints $\xi \geq 0$ and $\xi \leq 1$ by λ_0 and λ_1, respectively, the first-order condition for the optimal risky share ξ is given by

$$\mathbf{E}_i \left[\left(aR^p_{ih+1} + y_{h+1}\right)^{-\gamma} \left(R_{ih+1} - R_f\right) \right] + \lambda_0 - \lambda_1 = 0$$

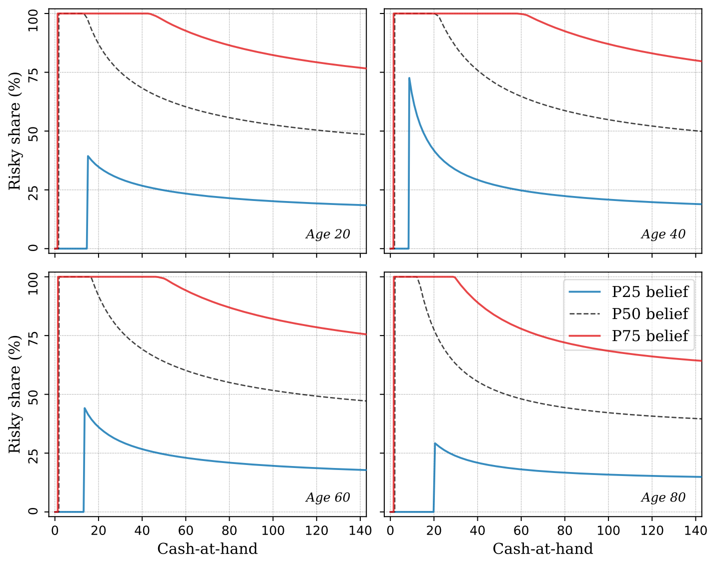


Figure 1.15: Household policy function for the optimal risky share. The red line shows the risky share adopted by an overly optimistic household (located at the 75th percentile of the belief distribution), while the dashed black line represents the choices of a household with median beliefs, and the blue line those of a pessimistic household at the 25th percentile.

where $\lambda_0 \geq 0$, $\lambda_1 \geq 0$ and the usual complementary-slackness conditions apply.

Consider a household that chooses to not participate, i.e., $\xi = 0$ which implies that $\lambda_0 > 0$ and $\lambda_1 = 0$: assuming that y_{h+1} is independent of R_{ih+1}, as is the case in this paper as well as in many others in the household-finance literature, the expectation can be split to obtain

$$\underbrace{\mathbf{E}_i\left[\left(aR_f + y_{h+1}\right)^{-\gamma}\right]}_{\text{expected MU}} \times \underbrace{\mathbf{E}_i\left[R_{ih+1} - R_f\right]}_{\text{risk premium}} + \lambda_0 = 0 \tag{1.13}$$

where the first term represents the expected marginal utility when saving everything in the risk-free asset, and the second term is the expected excess return (or risk premium). Expected marginal utility is unambiguously positive, and a rational expectations model usually imposes that $\mathbf{E}_i\left[R_{ih+1} - R_f\right] > 0$ in line with historical data on stock-market performance. In that case, the condition in (1.13) cannot hold for $\xi = 0$ as all terms on the l.h.s. are positive, and non-participation is therefore sub-optimal. On the other hand, in a model with subjective beliefs and learning from experience, households who had repeated low return realizations might, in fact, expect $\mathbf{E}_i\left[R_{ih+1} - R_f\right] < 0$, and thus (1.13) is satisfied for $\xi = 0$ and some $\lambda_0 > 0$.

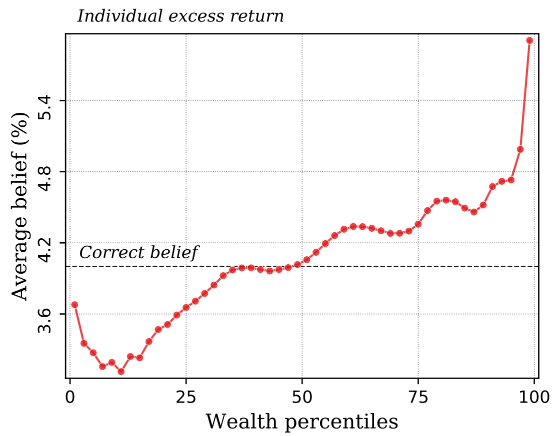


Figure 1.16: Average beliefs about individual excess returns along the wealth distribution. Each dot represents the average beliefs of households conditional on being in a given wealth percentile.

These households will choose not to hold the risky asset even in the absence of participation costs. As discussed above, due to the positive sorting across beliefs and wealth, such households will one average be poorer. This gradient of the perceived risk premium along the distribution is illustrated in Figure 1.16 for the benchmark model with participation costs.

1.8 Bayesian learning from experience (BLE)

The benchmark calibration uses the belief updating method from Malmendier and Nagel (2011, 2016). A natural alternative is to assume that agents incorporate new data in an optimal way using Bayes' rule, thus efficiently weighting all past observations, but ignoring information prior to becoming economically active themselves. The only difference to a standard rational-expectations model is thus the restriction that households only learn from their *own* experience.

Belief updating

In the Bayesian framework, households at any age h have a prior belief about mean excess returns, which I assume to be Gaussian and given by

$$\mu \sim \mathcal{N}\left(\widehat{\mu}_{ih}, \tau_h \sigma^2\right) \tag{1.14}$$

The prior is centered around $\widehat{\mu}_{ih}$, and the uncertainty associated with the belief is expressed as a scaling factor τ_h applied to the known variance of risky returns

σ^2. Within a cohort, there will be dispersion in households' mean belief $\widehat{\mu}_{ih}$ due to different realized return histories, while the variance $\tau_h \sigma^2$ is identical for all agents of the same age and is thus not indexed by i. This is a consequence of assuming that all newborns are assigned the same uncertainty about their beliefs τ_0 and update their beliefs each period, irrespective of whether they invest in stocks or not.

Newborns draw the initial beliefs about excess returns from a Gaussian distribution,

$$\widehat{\mu}_{i0} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}, \sigma_0^2\right) \tag{1.15}$$

which is centered around the true excess return. The variance σ_0^2 is set to the same value as in the benchmark model. Thus, both economies start off with the same distribution of mean beliefs.

As new return realizations are observed, households update their beliefs as follows: both the prior and the excess return were assumed to be Gaussian to obtain the usual analytically tractable case in which the posterior distribution is Gaussian and the updating rule is given by

$$\begin{aligned}\widehat{\mu}_{ih} &= \left[\frac{\left(\tau_h \sigma^2\right)^{-1}}{\left(\sigma^2\right)^{-1} + \left(\tau_h \sigma^2\right)^{-1}}\right] \widehat{\mu}_{ih-1} + \left[\frac{\left(\sigma^2\right)^{-1}}{\left(\sigma^2\right)^{-1} + \left(\tau_h \sigma^2\right)^{-1}}\right] \left(R_{ih} - R_f\right) \\ &= \frac{1}{1+\tau_h} \widehat{\mu}_{ih-1} + \frac{\tau_h}{1+\tau_h} \left(R_{ih} - R_f\right)\end{aligned}$$

Bayesian learning from experience therefore leaves the household problem virtually unchanged, except that the update weight α_h in (1.4) is now defined as

$$\alpha_h = \frac{\tau_h}{1+\tau_h}$$

The variance of the posterior distribution is given by the standard formula

$$\tau_{h+1} \sigma^2 = \frac{1}{\left(\tau_h \sigma^2\right)^{-1} + \left(\sigma^2\right)^{-1}}$$

which implies that the scaling factor evolves according to the non-linear difference equation

$$\tau_{h+1}^{-1} = \tau_h^{-1} + 1$$

Since $\lim_{h \to \infty} \tau_h = 0$, for infinitely-lived households their beliefs eventually collapse into a degenerate distribution around the true expected excess return, thus converging to the rational expectations equilibrium.

As illustrated in Figure 1.7, the weights assigned to past observations under the assumption of BLE differ from the setup in Malmendier and Nagel (2011, 2016). With BLE, the weight assigned to the most recent observation decreases more rapidly as

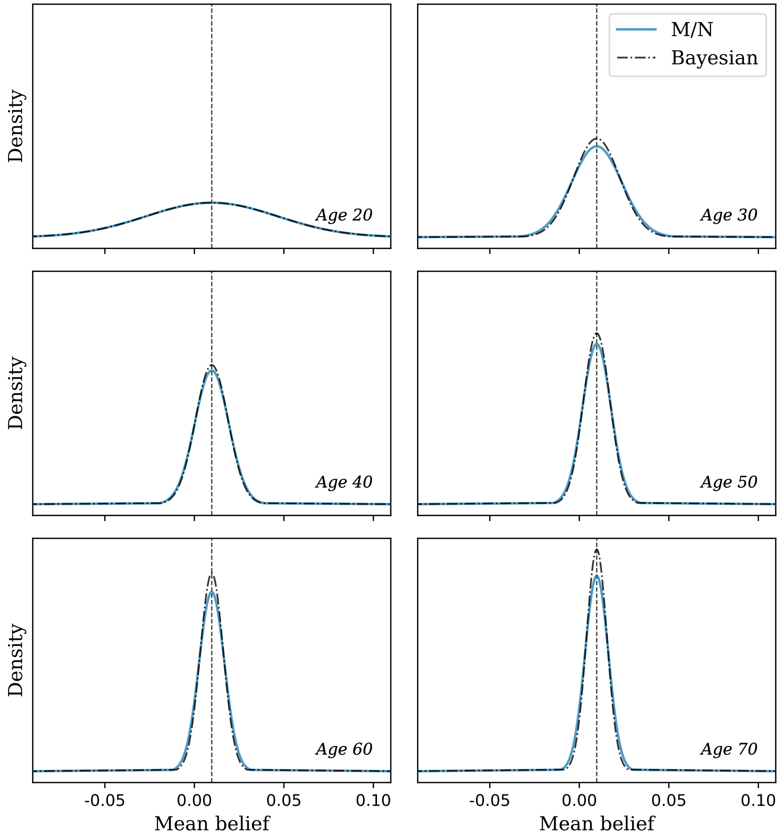


Figure 1.17: Cross-sectional distribution of mean beliefs at selected ages for Bayesian learning from experience and experience-based learning as in Malmendier and Nagel (2011, 2016).

households age, which implies that the *cross-sectional* distribution of mean beliefs clusters more tightly around the true excess return for older cohorts. This is illustrated in Figure 1.17 for selected ages: while the distributions are identical by construction at age 20 when newborns enter the economy, with Bayesian learning the cross-sectional dispersion collapses somewhat faster around the true expected excess return (which is $\approx 0.97\%$ in the quarterly calibration) as a cohort grows older. This, in turn, dampens any effect of subjective beliefs on the portfolio composition in the cross-section, moving the model closer to the rational-expectations framework.

Predictive distribution

Besides the faster convergence towards the true expected excess return, there is an additional conceptual difference as compared to the benchmark model: households factor in the uncertainty associated with their mean belief when evaluating the riskiness of their portfolio. For a household i of age h, the subjective expected risky return conditional on some belief μ is given by

$$\mathbf{E}_i\left[\, R_{ih+1} - R_f \,\middle|\, \mu \,\right] = \mathbf{E}_i\left[\, \mu + z_{ih+1} \,\middle|\, \mu \,\right] = \mu$$

and therefore, unconditionally, applying the law of iterated expectations, one obtains the same expression as in the Malmendier and Nagel (2011, 2016) setting, i.e.,

$$\mathbf{E}_i\left[\, R_{ih+1} - R_f \,\right] = \mathbf{E}_i\left[\, \mathbf{E}_i\left[\, R_{ih+1} - R_f \,\middle|\, \mu \,\right]\right] = \mathbf{E}_i\mu = \widehat{\mu}_{ih}$$

This no longer holds for the variance, however. With Bayesian learning, the conditional variance is

$$\mathrm{Var}_i\left(\, R_{ih+1} - R_f \,\middle|\, \mu \,\right) = \mathrm{Var}_i\left(\, \mu + z_{ih+1} \,\middle|\, \mu \,\right) = \sigma^2$$

while the unconditional variance, obtained from the law of total variance, evaluates to

$$\begin{aligned}\mathrm{Var}_i\left(\, R_{ih+1} - R_f \,\right) &= \mathbf{E}_i\left[\, \mathrm{Var}_i\left(\, R_{ih+1} - R_f \,\middle|\, \mu \,\right)\right] + \mathrm{Var}_i\left(\, \mathbf{E}_i\left[\, R_{ih+1} - R_f \,\middle|\, \mu \,\right]\right) \\ &= \sigma^2 + \tau_h\sigma^2\end{aligned}$$

The difference as compared to experience-based learning emerges because in the latter case, the beliefs are not uncertain from a household's perspective and thus, they do not amplify the perceived variance of risky returns. Conversely, the Bayesian updater acts as if the risky return process was effectively defined as

$$R_{ih+1} - R_f = \widehat{\mu}_{ih} + \sqrt{1 + \tau_h} \cdot z_{ih+1}\,, \qquad z_{ih+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma^2\right)$$

instead of (1.3). In principle, this additional perceived variance of risky returns stemming from the uncertainty of beliefs could be used to generate a lower participation or risky shares, in particular for young households with a higher τ_h. However, as I discuss in the next section, the effect turns out to be negligible.

Calibration

I use the same calibration as for the benchmark model. In particular, as mentioned above, I set the variance of initial mean beliefs σ_0^2 in (1.15) to the same value as in the benchmark listed in Table 1.2.

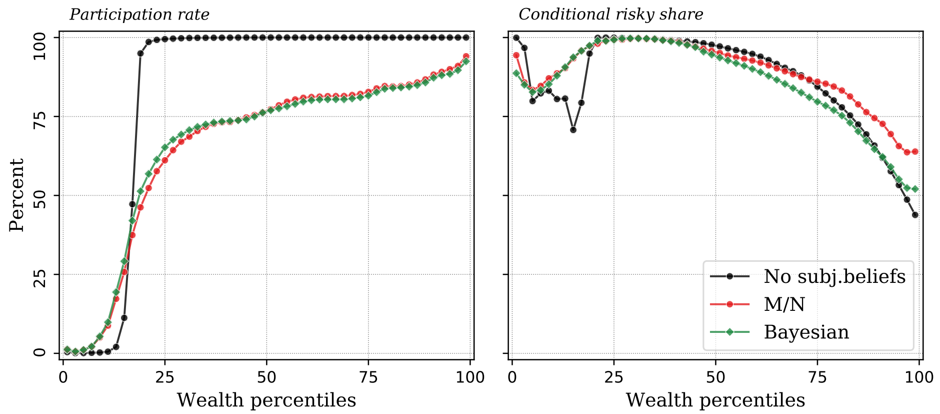


Figure 1.18: Portfolio composition along the wealth distribution. Bayesian learning from experience vs. benchmark.

The Bayesian learning model has one additional parameter τ_0 which controls newborn households' uncertainty about their beliefs. I set this parameter such that households at the age of 25 apply the same weight to the most recent realization as under experience-based learning, which is approximately 3.04% at a quarterly frequency. The initial scaling factor required to obtain the same value with Bayesian learning is $\tau_0 = 0.085$. Thus, newborns perceive the effective return variance to be only 8.5% higher than in the benchmark model, which is not sufficient to substantially affect their portfolio choices.

Results for the economy with Bayesian updating

The portfolio composition for the economy populated with Bayesian updaters is shown in Figure 1.18 and Figure 1.19 together with the benchmark model. Qualitatively, there are hardly any differences as compared to the benchmark model, neither across the wealth distribution, nor along the life-cycle. This result is hardly surprising given the preceding discussion, as the cross-sectional distributions of beliefs are close in both economies, and the perceived additional variance of risky returns due to belief uncertainty is modest. The results for the model *without* fixed participation costs (not shown) exhibit the same pattern, i.e., there are almost no differences between the two updating rules.

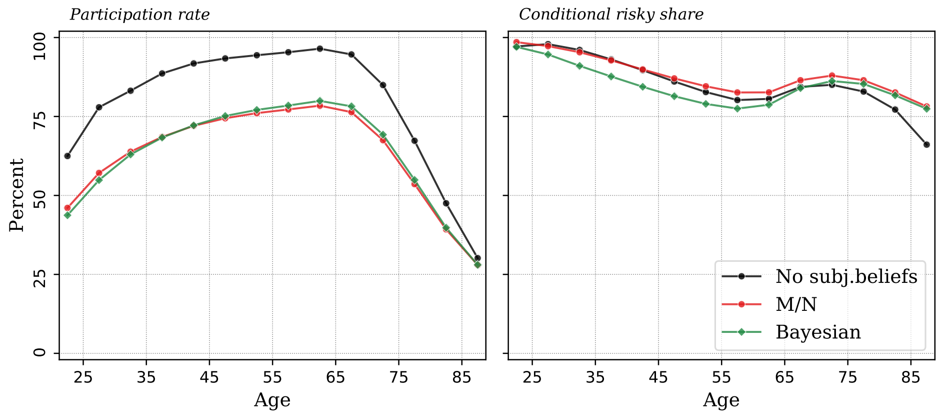


Figure 1.19: Portfolio composition over the life-cycle. Bayesian learning from experience vs. benchmark.

1.9 Correlated returns in the cross-section

In this section, I relax the assumption that returns are i.i.d. in the cross-section. An alternative model of household i's risky return is to assume a common component that arises because households invest a fraction of their portfolio into a market index. Denote by r^e_{mt} the excess return of such an index fund. Individual excess returns r^e_{it+1} in $t+1$ are then given by[19]

$$r^e_{it+1} = R_{it+1} - R_f = \beta_m r^e_{mt+1} + u_{it+1}$$

with

$$r^e_{mt+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}^m, \sigma^2_m\right) \qquad u_{it+1} \overset{\text{iid}}{\sim} \mathcal{N}\left(\overline{\mu}^u, \sigma^2_u\right) \qquad \operatorname{Corr}\left(r^e_{mt+1}, u_{it+1}\right) = 0$$

Here u_{it+1} is an uncorrelated idiosyncratic component that arises due to underdiversification. I assume that a household's portfolio choice is between investing in the risk-free asset and this composite risky asset, but households cannot choose the fraction of risky assets invested in the market index. The (objective) risk premium and the variance of this individual excess return are thus given by

$$\mathbf{E}\left[\, r^e_{it+1} \,\right] = \overline{\mu}^u + \beta_m \overline{\mu}^m \qquad \operatorname{Var}\left(\, r^e_{it+1} \,\right) = \beta^2_m \sigma^2_m + \sigma^2_u$$

Furthermore, the time-t correlation between any two households' gross returns is

$$\operatorname{Corr}\left(R_{it}, R_{jt}\right) = \frac{\beta^2_m \sigma^2_m}{\beta^2_m \sigma^2_m + \sigma^2_u}$$

19. Unlike in the previous sections where there was no need to distinguish between a household's age h and calendar time t, the notation in this section is adapted to reflect the fact that investors of different ages now receive the same market returns r^e_{mt} at time t.

I use the empirical counterparts of the above moments reported in Calvet, Campbell, and Sodini (2007) for Swedish data to pin down the additional parameters β_m, σ_m^2 and σ_u^2, as described in the calibration section below.

Implementing this model poses two additional challenges: first, households can be uncertain about both $\overline{\mu}^m$ and $\overline{\mu}^u$ and thus have to form beliefs about both, which introduces an additional continuous state variable into the household problem. Second, the economy no longer has a time-invariant ergodic distribution over the household's state variables, but depends on the sequence of aggregate realizations $(r_{mt}^e)_t$ which go back to the oldest cohort's date of birth.

In addition, this setting allows me to distinguish between public and private information. In the benchmark model it was assumed that households always observe the (potentially counterfactual) realization of their idiosyncratic return, regardless of whether they chose to invest in the risky asset or not. In this section, I instead assume that households always see the excess return on the market index r_{mt}^e which is public information. However, they observe the purely idiosyncratic return component u_{it} only if they choose to participate. This introduces a role for active learning, as households who pay the participation cost to invest in the stock market at the same time pay to acquire new information about $\overline{\mu}^u$ and can thus update their belief $\widehat{\mu}_{it}^u$.

Modified household problem

The household problem mostly remains the same as in the benchmark model of section 1.5, so I discuss the changes for the retired household only. Assuming that households are uncertain about the means of both the market and idiosyncratic returns, the expanded state vector is now given by $\boldsymbol{x} = \left(h, a, p, \widehat{\mu}_i^m, \widehat{\mu}_i^u, j\right)$ where $\widehat{\mu}_i^m$ is the belief about the market-return risk premium $\overline{\mu}^m$ and $\widehat{\mu}_i^u$ is the belief about $\overline{\mu}^u$. Retired households maximize

$$V_{jh}^r\left(a, p, \widehat{\mu}_i^m, \widehat{\mu}_i^u\right) = \max_{c, b, \xi} \left\{ c^{1-\psi} + \beta_j \left[\pi_h^s \mathbf{E}_i \left[\left(V_{jh+1}^r\left(a', p, (\widehat{\mu}_i^m)', (\widehat{\mu}_i^u)'\right)\right)^{1-\gamma} \right] + (1 - \pi_h^s)\mathbf{E}_i \left[\left(V_j^b\left(a_b'\right)\right)^{1-\gamma} \right] \right]^{\frac{1-\psi}{1-\gamma}} \right\}^{\frac{1}{1-\psi}}$$

subject to the same constraints as before. The belief about the market excess return is updated according to

$$(\widehat{\mu}_i^m)' = \left(1 - \alpha_{h+1}\right) \widehat{\mu}_i^m + \alpha_{h+1} r_{mt+1}^e$$

each period. On the other hand, $\widehat{\mu}_i^u$ is updated only if the household chooses a non-zero risky share ξ, i.e.,

$$(\widehat{\mu}_i^u)' = \begin{cases} (1-\alpha_{h+1})\widehat{\mu}_i^u + \alpha_{h+1} u_{it+1} & \text{if } \xi > 0 \\ \widehat{\mu}_i^u & \text{else} \end{cases}$$

Calibration

Calvet, Campbell, and Sodini (2007) report that in Swedish register data on household portfolios $\beta_m \approx 0.87$, which they obtain by regressing individual portfolio returns onto the MSCI world index. Additionally, the average idiosyncratic variance share in the Swedish data is reported to be

$$\frac{\sigma_u^2}{\beta_m^2 \sigma_m^2 + \sigma_u^2} \approx 60\% \,, \tag{1.16}$$

which implies that the cross-sectional correlation of returns is approximately 35%.[20]

To make the results as comparable to the benchmark model as possible, I choose the same overall risk premium and variance as reported in Table 1.2, i.e.,

$$\mathbf{E}_t \left[r_{it+1}^e \right] = \overline{\mu} \qquad\qquad \mathrm{Var}_t \left(r_{it+1}^e \right) = \sigma^2 \tag{1.17}$$

I assume that $\overline{\mu}^m = \overline{\mu}$ and $\sigma_m = \sigma$ since the values for $(\overline{\mu}, \sigma)$ imposed here are used in the household-finance literature to capture the moments of a broad market index in the first place. This implies that $\overline{\mu}^u = (1-\beta_m)\overline{\mu}$ needs to be imposed for (1.17) to hold. Finally, given $\beta_m = 0.87$, equation (1.16) pins the remaining parameter to be $\sigma_u = 17.05\%$ annually. This calibration is summarized in Table 1.4.

While one could alternatively assume that individual portfolios have different return moments than the market index, the above assumptions ensure that in a partial-equilibrium setting, a rational-expectations investor is indifferent to which share of the risky portfolio is invested in the market index, as the individual return moments are unaffected.

Lastly, in the model with market returns, the question arises which initial belief newborn investors should be assigned. Here I exactly follow the approach in Malmendier and Nagel (2011, 2016), who postulate that an individual of age $age_{it} = \underline{h} + h_{it}$ will form beliefs based on the returns observed in $t-1, \ldots t-\underline{h}-h_{it}$ going back to his or her year of birth. Thus, every newborn at time t has the same belief about market returns, as they all experienced the same history. As for the newborns' distribution of $\widehat{\mu}_{it}^u$, I adopt the same approach as in the benchmark model to generate an

20. Calvet, Campbell, and Sodini (2007) point out that both β_m and the idiosyncratic variance share are quite heterogeneous in the population. Given the already large state space of the model presented in this section, I ignore this additional source of heterogeneity.

Parameter	Description	Value	Source
Risky return			
$\overline{\mu}$	Risk premium	0.04	Cocco, Gomes, and Maenhout (2005)
σ	Volatility of risky return	0.16	Cocco, Gomes, and Maenhout (2005)
β_m	Market-β	0.87	Calvet, Campbell, and Sodini (2007)
–	Share of idiosyncratic variance	60%	Calvet, Campbell, and Sodini (2007)
Market return			
$\overline{\mu}^m$	Risk premium of market return	0.04	Cocco, Gomes, and Maenhout (2005)
σ_m	Volatility of market return	0.16	Cocco, Gomes, and Maenhout (2005)
Individual return			
$\overline{\mu}^u$	Risk premium of idiosyncratic return	$(1-\beta_m)\overline{\mu}$	–
σ_u	Volatility of idiosyncratic return	0.1705	–

Table 1.4: Parameters for the model with cross-sectionally correlated returns (annual). All remaining parameters are unchanged from the benchmark calibration.

initial distribution that is on average unbiased and has the cross-sectional standard deviation previously reported in Table 1.2.

Results

To interpret the results presented below, it is worthwhile to understand how these are generated. Since the economy now has an aggregate state given by the market return history going back to the birth date of the oldest cohort, every statistic used to characterize the portfolio composition in the cross-section or along the life-cycle depends on a particular realization of this market return time series. I address this issue by "bootstrapping" $N = 250$ such market return time series and computing the implied portfolio composition moments for each particular history.

Figure 1.20 shows the distribution of portfolio allocations along the wealth distribution from this bootstrapping exercise. Each dot represents the average of the 250 simulated economies for a given wealth bin, while the darker shaded areas represent the interquartile range, and the lighter shaded areas indicate the range which brackets 95% of the simulated economies.

This graph should be compared to the benchmark economy without aggregate market returns that was previously shown in Figure 1.12. Clearly the average of all simulated economies aligns well with both the participation rates and the conditional shares in that figure and thus, the findings from the benchmark model with i.i.d. returns continue to hold. Figure 1.21 shows the portfolio composition over the life-cycle and should be contrasted with Figure 1.13 for the benchmark model. Again, the mean over all simulations is quite close the the portfolio composition of the benchmark economy.

Both graphs show that the bootstrapped moments of the model without subjective

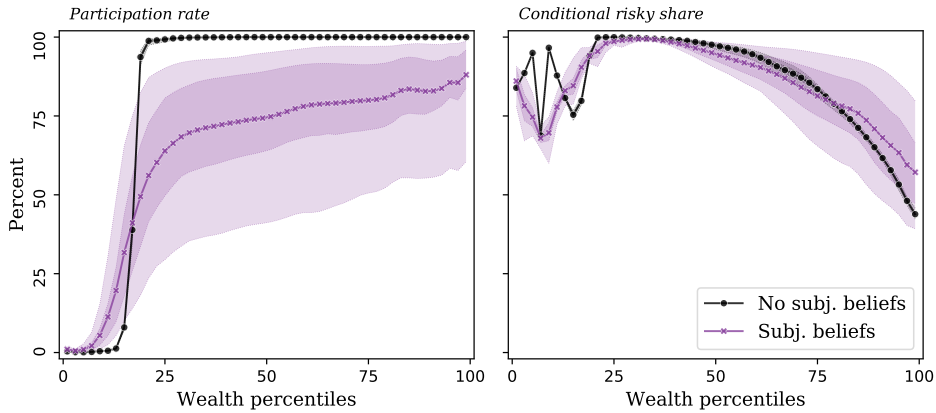


Figure 1.20: Portfolio composition along the wealth distribution. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Dark shaded areas show the interquartile range of simulated moments, while light shaded areas represent the range containing 95% of simulations.

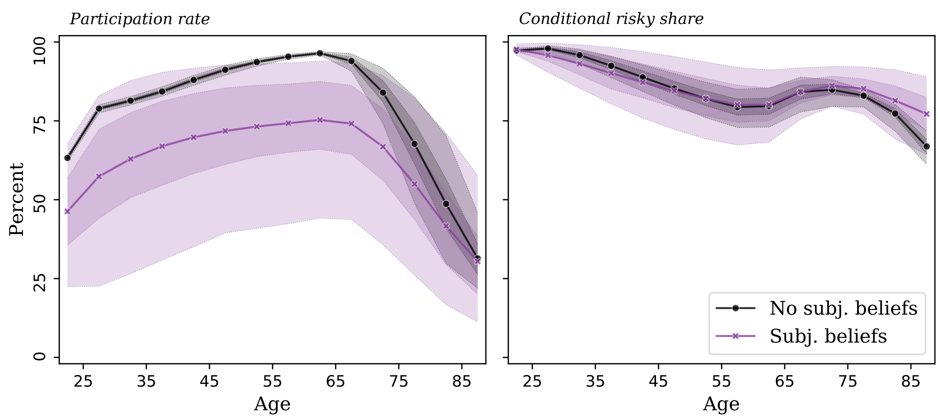


Figure 1.21: Portfolio composition over the life-cycle. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Dark shaded areas show the interquartile range of simulated moments, while light shaded areas represent the range containing 95% of simulations.

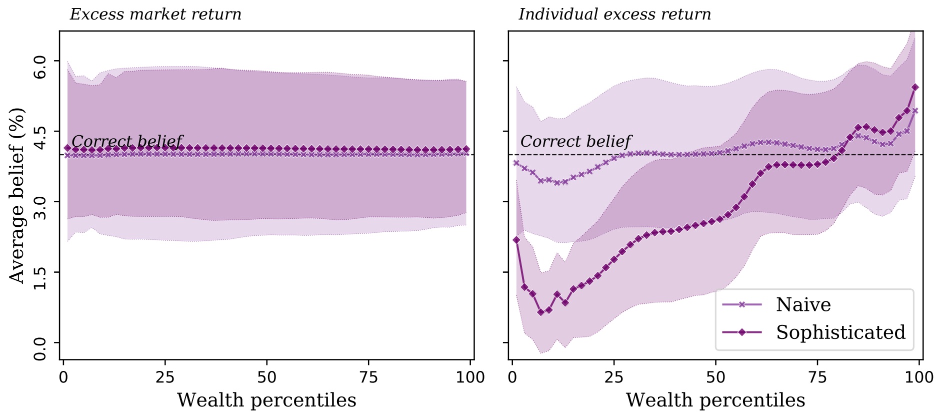


Figure 1.22: Beliefs about the excess market return r^e_{mt} and the individual excess return r^e_{it} for naive vs. sophisticated households in the model *with* participation costs. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

beliefs (shown in black) exhibit substantially less variation. In that economy, market return realizations (which are assumed to be i.i.d. over time) have almost no persistent effects other than on wealth accumulation, while in the model with subjective beliefs, they affect cohort-specific beliefs about market returns for a prolonged period of time.

In Figure 1.22, I report the analogue of Figure 1.16 for the economy with aggregate market returns. The beliefs about individual mean excess returns shown in Figure 1.22 are computed as $\widehat{\mu}_{it} = \beta_m \widehat{\mu}^m_{it} + \widehat{\mu}^u_{it}$. The graph illustrates that just like in the benchmark model, there is a positive sorting across beliefs about *individual* returns and wealth (right-hand panel); however, on average no such sorting exists with respect to beliefs about *market* returns (left-hand panel). This is due to the fact that conditional on calendar year and age, all households hold the same beliefs about the market return irrespective of their wealth.

The right-hand panel of Figure 1.22 also highlights that in the setting discussed in this section, the differences between economies populated by naive vs. sophisticated agents are more pronounced. Conversely, in the benchmark model, where households update their beliefs even if they do not participate, naive and sophisticated households hold almost identical beliefs, as shown in Figure 1.39 in the appendix.

If belief updating is conditional on participation, and participation is costly due to per-period participation costs, the beliefs held by naive and sophisticated agents diverge. The reason is the following: sophisticated households are in a sense more risk-averse because they anticipate that in case of a low return realization tomorrow,

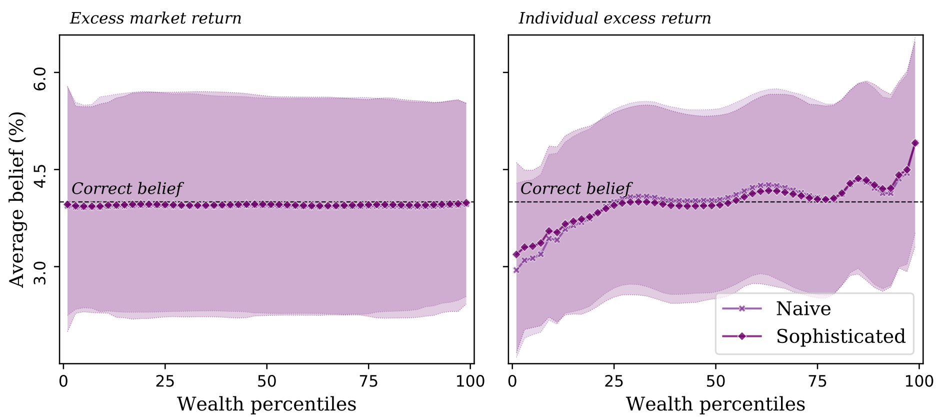


Figure 1.23: Beliefs about the excess market return r^e_{mt} and the individual excess return r^e_{it} for naive vs. sophisticated households in the model *without* participation costs. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

they will adjust their beliefs about excess returns downward. Thus, they expect to earn a lower return on the investment choices they make tomorrow. Given that the EIS is assumed to be lower than one, they anticipate to *increase* their savings tomorrow when a bad shock hits, thereby lowering consumption. Low returns tomorrow consequently coincide with higher marginal utility tomorrow, which tilts the risky share chosen *today* downward (this is evident from the policy functions shown in Figure 1.40 in the appendix for the benchmark model). In the presence of participation costs, households who would choose a lower risky share conditional on participating have a higher incentive not to hold the risky asset at all. This effect is even more pronounced for households who are already more pessimistic about (idiosyncratic) returns, and since they choose not to participate in this case, they never update their pessimistic beliefs about idiosyncratic returns. Via this mechanism, sophisticated households are more likely to end up in non-participation as an absorbing state. Figure 1.23 shows that costly participation is a key for this mechanism to be operational: without it, there is almost no difference between the cross-sectional beliefs in an economy with naive vs. one with sophisticated agents.

Lastly, Figure 1.24 and Figure 1.25 contrast the portfolio composition across the wealth distribution and over the life-cycle for naive vs. sophisticated agents. These graphs should be compared to those for the benchmark economy shown in Figure 1.37 and Figure 1.38 in the appendix. For the reasons discussed above, the differences in the present model compared to the benchmark are now more substantial, also among middle-class and younger households.

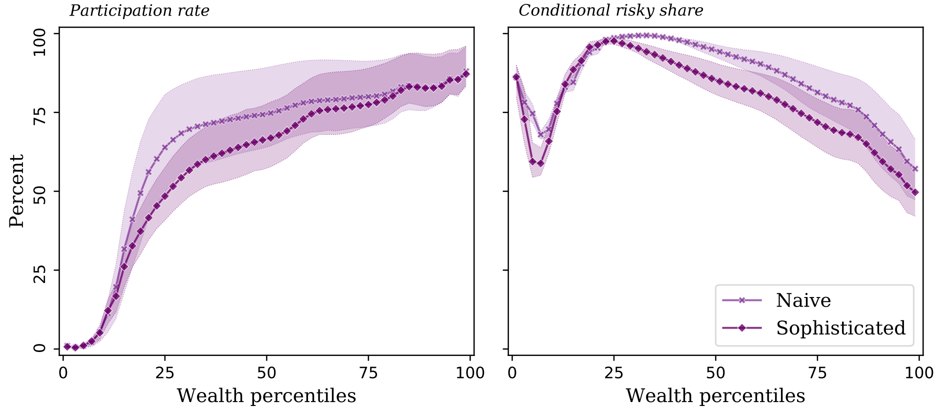


Figure 1.24: Portfolio composition along the wealth distribution. Naive vs. sophisticated households. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

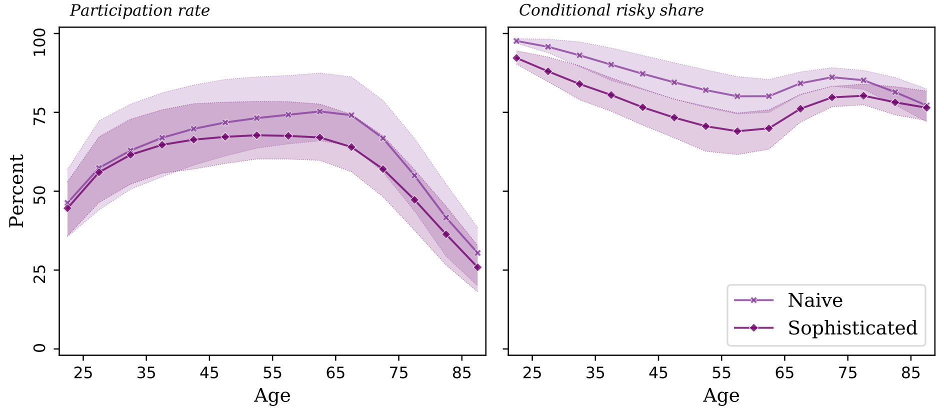


Figure 1.25: Portfolio composition over the life-cycle. Naive vs. sophisticated households. Model *with* participation costs and aggregate market returns. Dots indicate bin averages over simulated economies. Shaded areas show the interquartile range of simulated moments.

1.10 Summary and conclusion

In this paper, I propose a mechanism to help explain the empirically observed heterogeneity of households' financial portfolios along the wealth distribution, in particular the almost monotonically increasing stock market participation rate. To this end, I incorporate empirical evidence on how households form beliefs based on past realizations of asset returns into an otherwise standard household-finance model. Unlike in a fully rational model, households overweight past observations of returns on their idiosyncratic portfolios when forming beliefs about excess returns, thus generating a belief dispersion in the cross-section that does not fully disappear over time – in this setting, households do not necessarily learn the true excess return in their lifetime.

Combined with underdiversification and thus idiosyncratic return histories, experience-based learning of this kind introduces heterogeneity in households' participation and portfolio allocation choices in a way that moves the model's predictions closer to the data. As above-average return realizations simultaneously move a household upwards in the wealth distribution and improve the household's outlook on future returns, this mechanism induces positive sorting in the joint distribution of wealth and beliefs: on average, the most wealthy households end up being the most optimistic and, consequently, they increase the fraction held in risky assets compared to a fully-rational agent. The opposite holds for households at the other end of the wealth distribution, at least to the extent that they are poor due to low return realizations in the past, and the participation rate within this group drops as compared to a standard model.

It is a quantitative question of how much of this intuitively appealing mechanism survives in a full life-cycle model where additional factors affect a household's wealth but not its beliefs about stock returns. With a highly-persistent earnings uncertainty and a hump-shaped wealth profile over the life-cycle, asset returns are neither the only nor the most dominant driver of a household's position in the wealth distribution. The wealthy in this framework are rich because they were lucky in how their investment paid off, or because they had repeated high earnings draws – the latter, however, have no effect on their beliefs about excess returns, and hence on their portfolio choice. Additionally, in a life-cycle framework, the newborns enter the economy without a history of past returns correlated with their position in the wealth distribution, further muting the effect of subjective beliefs.

The preceding sections have shown that subjective beliefs and experience-based learning are able to generate limited participation even in such a setting. Adding a reasonably small participation cost further improves the model fit with the data, above and beyond what a standard model with the same participation cost can achieve, in particular when it comes to the limited participation observed among middle-class

households. This indicates that the mechanism presented here can potentially interact with other building blocks to move the model prediction even closer to the data. Exploring such extensions is left to future research.

Bibliography

Ameriks, John, Gábor Kézdi, Minjoon Lee, and Matthew D. Shapiro. 2019. Heterogeneity in Expectations, Risk Tolerance, and Household Stock Shares: The Attenuation Puzzle. *Journal of Business & Economic Statistics* 0 (0): 1–27.

Arias, Elizabeth, and Jianquan Xu. 2018. United States life tables, 2015. *National vital statistics reports,* Working Paper Series, 67 (2).

Arrondel, Luc, Calvo-Pardo Hector, and Derya Tas. 2014. *Subjective return expectations, information and stock market participation: evidence from France.* Discussion Paper Series In Economics And Econometrics. University of Southampton, Economics Division, School of Social Sciences.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2018. *Rich Pickings? Risk, Return, and Skill in the Portfolios of the Wealthy.* Technical report.

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2019. Wealth Distribution and Social Mobility in the US: A Quantitative Approach. *American Economic Review* 109 (5): 1623–1647.

Blanchard, Olivier J. 1985. Debt, Deficits, and Finite Horizons. *Journal of Political Economy* 93 (2): 223–247.

Bommier, Antoine, Daniel Harenberg, and François Le Grand. 2017. Recursive Preferences, the Value of Life, and Household Finance.

Briggs, Joseph S., David Cesarini, Erik Lindqvist, and Robert Östling. 2019. *Windfall Gains and Stock Market Participation.* Technical report. National Bureau of Economic Research, Inc.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. 2007. Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy* 115 (5): 707–747.

Campanale, Claudio, Carolina Fugazza, and Francisco Gomes. 2015. Life-cycle portfolio choice with liquid and illiquid financial assets. *Journal of Monetary Economics* 71:67–83.

Campbell, John Y. 2006. Household Finance. *The Journal of Finance* 61 (4): 1553–1604.

Catherine, Sylvain. 2019. *Countercyclical Labor Income Risk and Portfolio Choices over the Life-Cycle.*

Chang, Yongsung, Jay H. Hong, and Marios Karabarbounis. 2018. Labor Market Uncertainty and Portfolio Choice Puzzles. *American Economic Journal: Macroeconomics* 10 (2): 222–262.

Choi, James J., David Laibson, Brigitte C. Madrian, and Andrew Metrick. 2009. Reinforcement Learning and Savings Behavior. *Journal of Finance* 64 (6): 2515–2534.

Cocco, João F. 2004. Portfolio Choice in the Presence of Housing. *The Review of Financial Studies* 18 (2): 535–567.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout. 2005. Consumption and Portfolio Choice over the Life Cycle. *Review of Financial Studies* 18 (2): 491–533.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016a. Asset Pricing When 'This Time Is Different'. *The Review of Financial Studies* 30 (2): 505–535.

———. 2016b. Parameter Learning in General Equilibrium: The Asset Pricing Implications. *American Economic Review* 106 (3): 664–698.

De Nardi, Mariacristina. 2004. Wealth Inequality and Intergenerational Links. *Review of Economic Studies* 71 (3): 743–768.

Dominitz, Jeff, and Charles F. Manski. 2007. Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study. *Journal of the European Economic Association* 5 (2-3): 369–379.

———. 2011. Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics* 26 (3): 352–370.

Druedahl, Jeppe, and Thomas Høgholm Jørgensen. 2017. A general endogenous grid method for multi-dimensional models with non-convexities and constraints. *Journal of Economic Dynamics and Control* 74:87–107.

Ehling, Paul, Alessandro Graniero, and Christian Heyerdahl-Larsen. 2017. Asset Prices and Portfolio Choice with Learning from Experience. *The Review of Economic Studies* 85 (3): 1752–1780.

Epstein, Larry G. 1988. Risk aversion and asset prices. *Journal of Monetary Economics* 22 (2): 179–192.

Epstein, Larry G., and Stanley E. Zin. 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57 (4): 937–969.

Fagereng, Andreas, Charles Gottlieb, and Luigi Guiso. 2017. Asset Market Participation and Portfolio Choice over the Life-Cycle. *Journal of Finance* 72 (2): 705–750.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2019. *Heterogeneity and Persistence in Returns to Wealth.* Technical report. National Bureau of Economic Research, Inc.

Foltyn, Richard, and Jonna Olsson. 2019. *Subjective life expectancies, time preference heterogeneity and wealth inequality.*

Gaudecker, Hans-Martin von, and Axel Wogrolly. 2019. The dynamics of households' stock market beliefs.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2019. *Five Facts About Beliefs and Portfolios.* Working Paper, Working Paper Series 25744. National Bureau of Economic Research.

Gomes, Francisco, and Alexander Michaelides. 2003. Portfolio choice with internal habit formation: a life-cycle model with uninsurable labor income risk. Finance and the Macroeconomy, *Review of Economic Dynamics* 6 (4): 729–766.

———. 2005. Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence. *Journal of Finance* 60 (2): 869–904.

Greenwood, Robin, and Andrei Shleifer. 2014. Expectations of Returns and Expected Returns. *The Review of Financial Studies* 27 (3): 714–746.

Guiso, Luigi, and Paolo Sodini. 2013. Household Finance: An Emerging Field. Chap. 21, edited by George M. Constantinides, Milton Harris, and Rene M. Stulz, 2:1397–1532. Handbook of the Economics of Finance. Elsevier.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2016. *What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk?* Technical report.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. The Nature of Countercyclical Income Risk. *Journal of Political Economy* 122 (3): 621–660.

Haliassos, Michael, and Carol C. Bertaut. 1995. Why do so Few Hold Stocks? *The Economic Journal* 105 (432): 1110–1129.

Heiss, Florian, Michael Hurd, Maarten van Rooij, Tobias Rossmann, and Joachim K Winter. 2019. Dynamics and heterogeneity of subjective stock market expectations.

Hendricks, Lutz. 2007. How important is discount rate heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control* 31 (9): 3042–3068.

Hubmer, Joachim, Per Krusell, and Anthony A. Jr. Smith. 2019. *Sources of U.S. Wealth Inequality: Past, Present, and Future.* Technical report.

Hurd, Michael D, and Susann Rohwedder. 2012. *Stock Price Expectations and Stock Trading.* Working Paper, Working Paper Series 17973. National Bureau of Economic Research.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning. 2017. The endogenous grid method for discrete-continuous dynamic choice models with (or without) taste shocks. *Quantitative Economics* 8 (2): 317–365.

Kaustia, Markku, and Samuli Knüpfer. 2008. Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions. *Journal of Finance* 63 (6): 2679–2702.

Kézdi, Gábor, and Robert J Willis. 2011. *Household Stock Market Beliefs and Learning.* Working Paper, Working Paper Series 17614. National Bureau of Economic Research.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2015. *The Tail that Wags the Economy: Beliefs and Persistent Stagnation.* Working Paper, Working Paper Series 21719. National Bureau of Economic Research.

Krueger, Dirk, Kurt Mitman, and Fabrizzio Perri. 2016. Macroeconomics and Household Heterogeneity, edited by John B. Taylor and Harald Uhlig, 2:843–921. Handbook of Macroeconomics. Elsevier.

Krusell, Per, and Anthony A. Smith. 1997. Income And Wealth Heterogeneity, Portfolio Choice, And Equilibrium Asset Returns. *Macroeconomic Dynamics* 1 (02): 387–422.

———. 1998. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy* 106 (5): 867–896.

Kuhn, Moritz, Moritz Schularick, and Ulrike I Steins. 2019. *Income and Wealth Inequality in America.*

Malmendier, Ulrike, and Stefan Nagel. 2011. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *The Quarterly Journal of Economics* 126 (1): 373–416.

———. 2016. Learning from Inflation Experiences. *The Quarterly Journal of Economics* 131 (1): 53–87.

Malmendier, Ulrike, Demian Pouzo, and Victoria Vanasco. 2018. *Investor Experiences and Financial Market Dynamics.* NBER Working Papers 24697. National Bureau of Economic Research, Inc.

Mankiw, N.Gregory. 1986. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics* 17 (1): 211–219.

Merton, Robert C. 1971. Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory* 3 (4): 373–413.

Meyer, Steffen, and Michaela Pagel. 2019. *Fully Closed: Individual Responses to Realized Gains and Losses.* Working Paper, Working Paper Series 25542. National Bureau of Economic Research.

Nagel, Stefan, and Zhengyang Xu. 2019. Asset Pricing with Fading Memory.

Nakov, Anton, and Galo Nuño. 2015. Learning from experience in the stock market. *Journal of Economic Dynamics and Control* 52:224–239.

Poterba, James M., and Andrew A. Samwick. 1995. Stock Ownership Patterns, Stock Market Fluctuations, and Consumption. *Brookings Papers on Economic Activity* 1995 (2).

Schraeder, Stefanie. 2015. Information Processing and Non-Bayesian Learning in Financial Markets. *Review of Finance* 20 (2): 823–853.

Storesletten, Kjetil, Chris Telmer, and Amir Yaron. 2007. Asset Pricing with Idiosyncratic Risk and Overlapping Generations. *Review of Economic Dynamics* 10 (4): 519–548.

Vestman, Roine. 2018. Limited Stock Market Participation Among Renters and Homeowners. *Review of Financial Studies* 32 (4): 1494–1535.

Vissing-Jorgensen, Annette. 2003. Perspectives on Behavioral Finance: Does ”Irrationality”Disappear with Wealth? Evidence from Expectations and Actions. *NBER Macroeconomics Annual* 18:139–194.

Wachter, Jessica A., and Motohiro Yogo. 2010. Why Do Household Portfolio Shares Rise in Wealth? *The Review of Financial Studies* 23 (11): 3929–3965.

Weil, Philippe. 1990. Nonexpected Utility in Macroeconomics. *The Quarterly Journal of Economics* 105 (1): 29–42.

Yao, Rui, and Harold H. Zhang. 2004. Optimal Consumption and Portfolio Choices with Risky Housing and Borrowing Constraints. *The Review of Financial Studies* 18 (1): 197–239.

Appendix

1.A Additional evidence from the Survey of Consumer Finances

Detailed household balance sheets

	Avg. (in \$)	Median (in \$)	Participation (in %)
Checking accounts	4,769	1,211	82.1
Savings accounts	10,243	105	55.0
Life insurance	9,131	0	26.0
Mutual funds (safe)	7,454	0	5.8
Retirement accounts (safe)	31,084	0	43.0
Total safe assets	98,638	14,149	92.9
Stocks	41,262	0	20.0
Mutual funds (risky)	20,749	0	14.2
Retirement accounts	44,957	0	46.2
Total risky assets	128,682	1,514	54.8
Total financial assets	227,320	24,631	93.1
Consumer debt	3,000	0	47.4
Education loans	2,633	0	13.7
Home ownership	–	–	69.4
Housing wealth (owners)	344,354	181,109	100.0
Mortgages (owners)	106,355	59,057	70.2
Net housing wealth (owners)	237,999	101,346	100.0
Actively managed businesses	89,669	0	11.3
Passively-held businesses	10,439	0	1.4
Total gross wealth	555,657	162,606	94.8
Net worth	477,458	93,742	–

Table 1.5: Summary statistics for disaggregated households' balance sheets. Housing wealth and mortgage statistics are reported for the sub-sample of homeowners. Data source: SCF 1998–2007.

	Avg. (in \$)	Median (in \$)	Participation (in %)
Checking accounts	5,122	1,333	88.2
Savings accounts	11,002	263	59.1
Life insurance	9,807	0	27.9
Mutual funds (safe)	8,006	0	6.2
Retirement accounts (safe)	33,376	0	46.2
Total safe assets	105,940	17,697	99.8
Stocks	44,316	0	21.5
Mutual funds (risky)	22,285	0	15.2
Retirement accounts	48,272	0	49.6
Total risky assets	138,208	3,356	58.8
Total financial assets	244,148	31,254	100.0
Consumer debt	3,196	0	49.8
Education loans	2,781	0	14.1
Home ownership	–	–	72.8
Housing wealth (owners)	350,836	182,850	100.0
Mortgages (owners)	108,341	61,060	70.7
Net housing wealth (owners)	242,495	103,491	100.0
Actively managed businesses	96,255	0	12.0
Passively-held businesses	11,212	0	1.5
Total gross wealth	595,420	186,239	100.0
Net worth	511,933	111,811	–

Table 1.6: Summary statistics for disaggregated households' balance sheets for sub-sample with *positive* financial wealth. Housing wealth and mortgage statistics are reported for the sub-sample of homeowners. Data source: SCF 1998–2007.

Portfolio composition across the wealth distribution

In this section, I present additional evidence on the composition of household portfolios observed in the SCF. In Figure 1.26, I plot net total wealth (net worth) on the x-axis, which includes all wealth components (including housing) after deducting all household liabilities (including mortgages). Figure 1.27 shows the portfolio allocation along percentiles of gross financial wealth, while Figure 1.28 reports portfolio choices over the distribution of financial wealth net credit-card debt and consumer loans. Last, Figure 1.29 plots those moments after additionally subtracting education loans from safe financial assets.

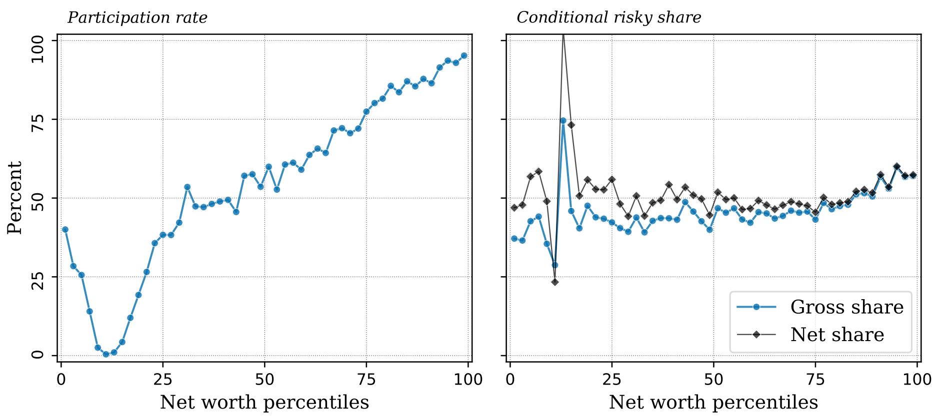


Figure 1.26: Portfolio composition along percentiles of *net worth*. Each dot represents two percent of households. Data source: SCF 1998–2007.

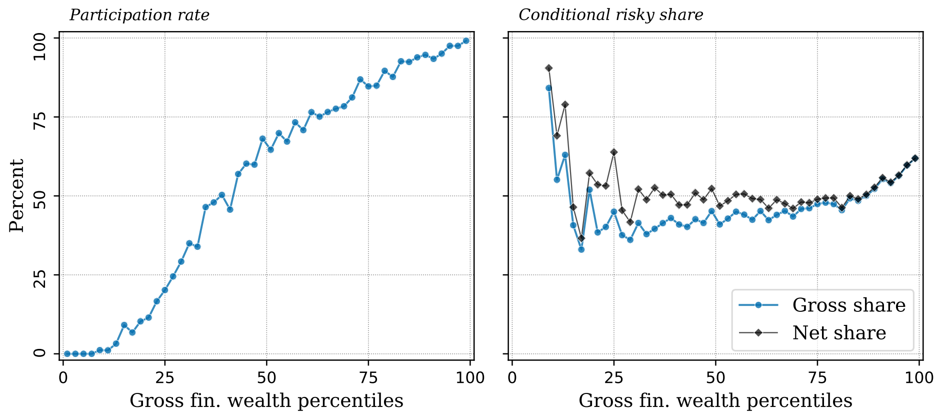


Figure 1.27: Portfolio composition along percentiles of *gross financial wealth*. Each dot represents two percent of households. Data source: SCF 1998–2007.

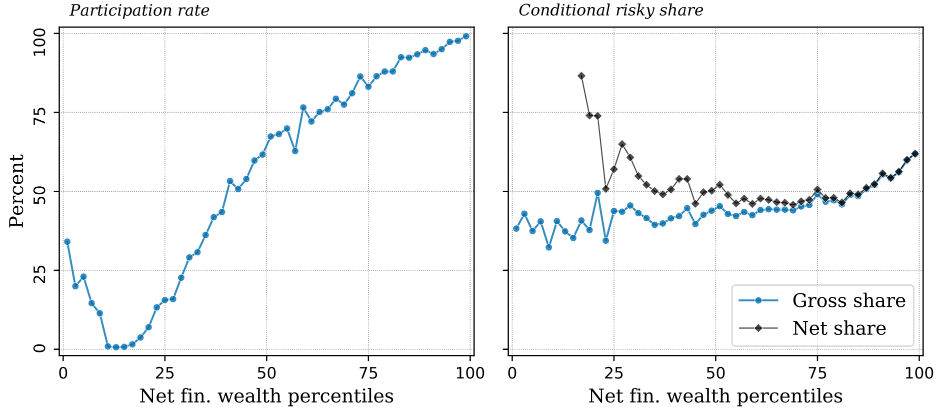


Figure 1.28: Portfolio composition along percentiles of *financial wealth net of consumer debt*. Each dot represents two percent of households. Data source: SCF 1998–2007.

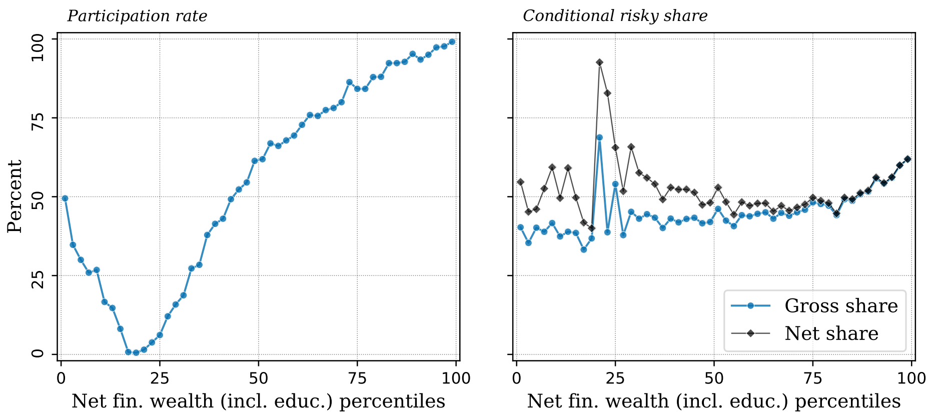


Figure 1.29: Portfolio composition along percentiles of *financial wealth net of consumer debt and education loans*. Each dot represents two percent of households. Data source: SCF 1998–2007.

Disaggregation by homeownership status

This section contains additional graphs disaggregating household portfolio allocations by homeownership status. In Figure 1.30, I show that financial portfolios are comparable conditioning on net worth. Figure 1.31 reports portfolio choices by gross financial wealth, while Figure 1.32 and Figure 1.33 illustrate that the differences are even smaller when plotting against the deciles of net financial wealth.

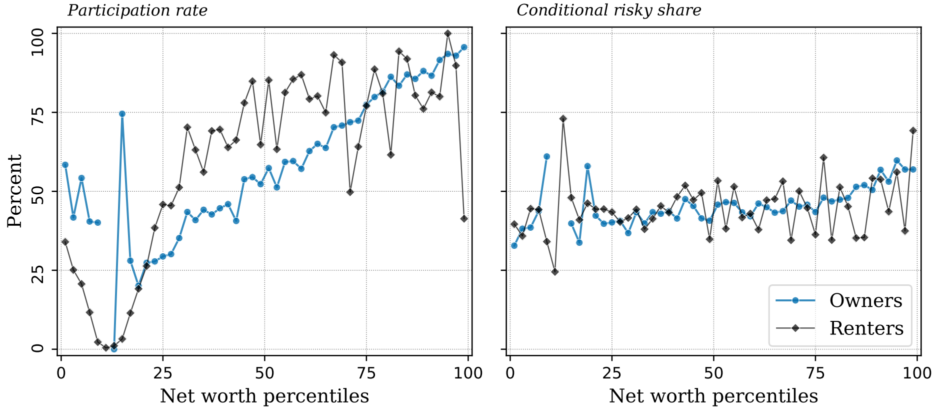


Figure 1.30: Portfolio composition along percentiles of *net worth* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

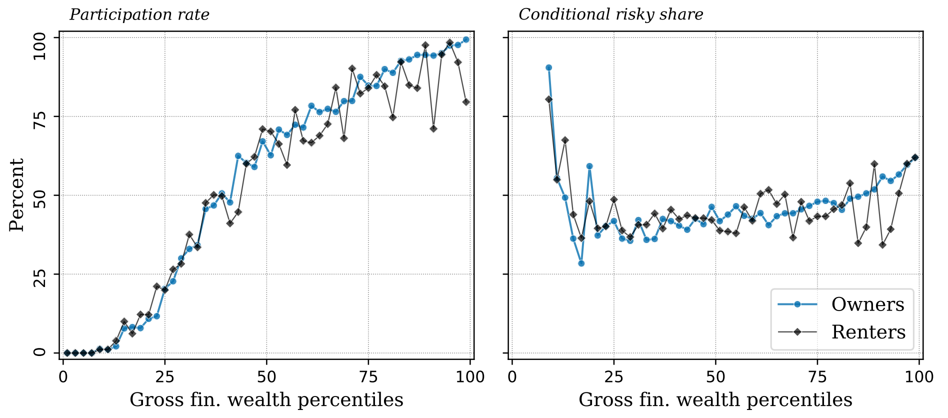


Figure 1.31: Portfolio composition along percentiles of *gross financial wealth* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

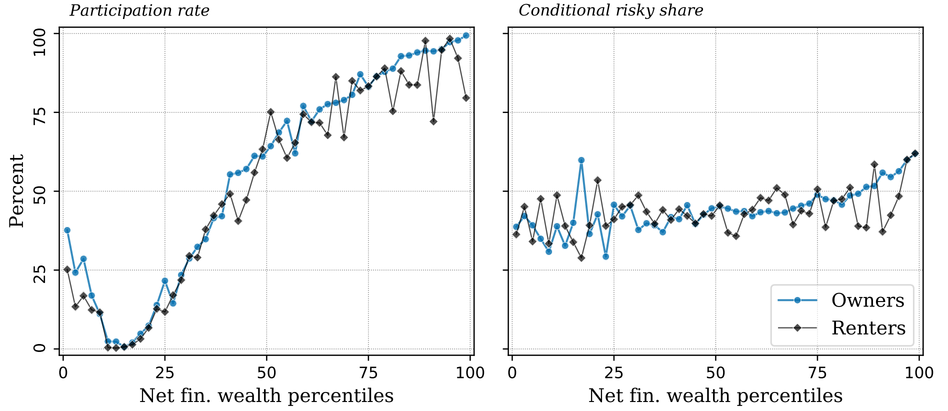


Figure 1.32: Portfolio composition along percentiles of *financial wealth net of consumer debt* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

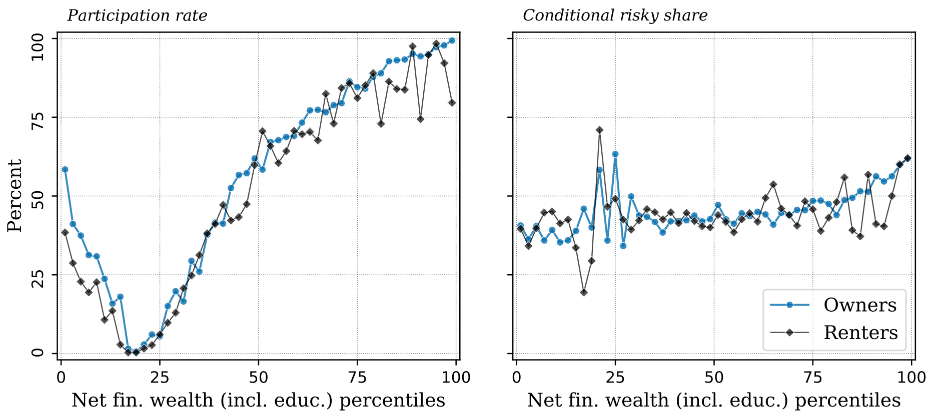


Figure 1.33: Portfolio composition along percentiles of *financial wealth net of consumer debt and education loans* by homeownership status. Wealth percentiles are computed for the *pooled* sample of owners and renters. Plots show averages conditional on wealth percentile and homeownership status. Data source: SCF 1998–2007.

1.B Recursive formulation of experience-based learning

In their paper, Malmendier and Nagel (2011) postulate that the beliefs about returns in period t are a weighted average of past realizations prior to t, going back as far as a person's year of birth. The index of historical returns proposed by Malmendier and Nagel (2011) is given by equations (1.5) and (1.6) in the main text.

Denote the new excess return observation at the beginning of period t by $r^e_{it} = R_t - R_f$. In this section, I ignore whether R_t has an idiosyncratic return component and drop any household-specific index i. The historical return index in Malmendier and Nagel (2011), adjusted to the notation used in the present paper and the fact that age $h = 0$ in the model corresponds to $\underline{h}$ actual years, can then be written as

$$\widehat{\mu}_{it} = \frac{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}\, r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}$$

where $\widehat{\mu}_{it}$ is the belief about mean excess returns by agent i of age h_{it} in period t after observing the risky realization at the beginning of period t. Note that, as in Malmendier and Nagel (2011), the above expression is not defined for $\underline{h} + h_{it} \leq 1$, i.e., they assume that the first return realization taken into account is the one that occurred at age one.

The recursive updating weight can be derived as follows:

$$\begin{aligned}
\widehat{\mu}_{it} &= \frac{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda} r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t + \frac{\sum_{k=2}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda} r^e_{t-k+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t + \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda} r^e_{t-(k-1)+1}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \\
&= \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} r^e_t \\
&\quad + \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} \frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda} r^e_{t-(k-1)+1}}{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}} \\
&= \left[\frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}\right] r^e_t + \left[\frac{\sum_{k=1}^{\underline{h}+(h_{it}-1)-1} (\underline{h} + (h_{it} - 1) - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}\right] \widehat{\mu}_{it-1} \\
&= \left[\frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}\right] r^e_t + \left[\frac{\sum_{k=2}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}\right] \widehat{\mu}_{it-1}
\end{aligned}$$

Furthermore, since

$$1 - \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}} = \frac{\sum_{k=2}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}$$

the recursive updating equation can be written as

$$\widehat{\mu}_{it} = (1 - \alpha_{it})\widehat{\mu}_{it-1} + \alpha_{it} r_t^e$$

with the weight on the most recent observation given by

$$\alpha_{it} = \frac{(\underline{h} + h_{it} - 1)^{\lambda}}{\sum_{k=1}^{\underline{h}+h_{it}-1} (\underline{h} + h_{it} - k)^{\lambda}}$$

1.C Alternative calibration similar to Cocco et al. (2005)

This section illustrates how the model with subjective beliefs performs using a more "standard" calibration. To this end, to the extent possible, I adopt the parameters used in Cocco, Gomes, and Maenhout (2005), which have become somewhat of a benchmark model in the household-finance literature, and contrast the resulting portfolio allocation across the wealth distribution and over the life-cycle.[21] Unlike in the rest of the paper, I therefore set the risk-aversion to $\gamma = 10$, the elasticity of intertemporal substitution to $1/\gamma$, the annual discount factor to $\beta = 0.96$, and turn off the bequest motive. As in Cocco, Gomes, and Maenhout (2005), the participation cost κ is set to zero and there is no preference heterogeneity, unlike in the main text. These parameters are summarized in Table 1.7.

Parameter	Description	Value	Source
β	Discount factor (annualized)	0.96	Cocco, Gomes, and Maenhout (2005)
γ	Risk aversion	10	Cocco, Gomes, and Maenhout (2005)
ψ^{-1}	Elasticity of intertemporal substitution	0.1	Cocco, Gomes, and Maenhout (2005)
ϕ	Weight on bequest utility	0	Cocco, Gomes, and Maenhout (2005)
κ	Participation cost	0	Cocco, Gomes, and Maenhout (2005)

Table 1.7: Parameters used for alternative calibration, taken from Cocco, Gomes, and Maenhout (2005) where applicable. The remaining parameters are unchanged from the benchmark calibration in the main text.

The resulting portfolio composition across the wealth distribution is shown in Figure 1.34, and Figure 1.35 plots averaged household choices over the life-cycle.

Whereas participation rates along the wealth distribution look similar to the results reported for the model without participation costs in the main text, the conditional risky share in this calibration is considerably lower. The reason is that this "standard" calibration performs poorly in terms of matching the wealth distribution, generating households that hold substantially more assets than in the data. Since the optimal risky share is decreasing in wealth, this generates a lower conditional risky share in the cross-section.

To illustrate this, I plot gross total wealth as reported in the SCF for each wealth decile (approximately \$117,000 for the 5th decile), expressed as multiples of average quarterly earnings, and contrast these values with their model counterparts.[22] Evi-

21. My model implementation does not have any heterogeneous income profiles for three education groups as in Cocco, Gomes, and Maenhout (2005). Additionally, their stochastic earnings process features a unit root, while the one used here follows an AR(1), albeit with a very high persistence. The results for the "standard" model presented here are thus not identical to those in Cocco, Gomes, and Maenhout (2005).

22. I omit the highest decile as this makes the differences for the lower deciles hard to read. Both

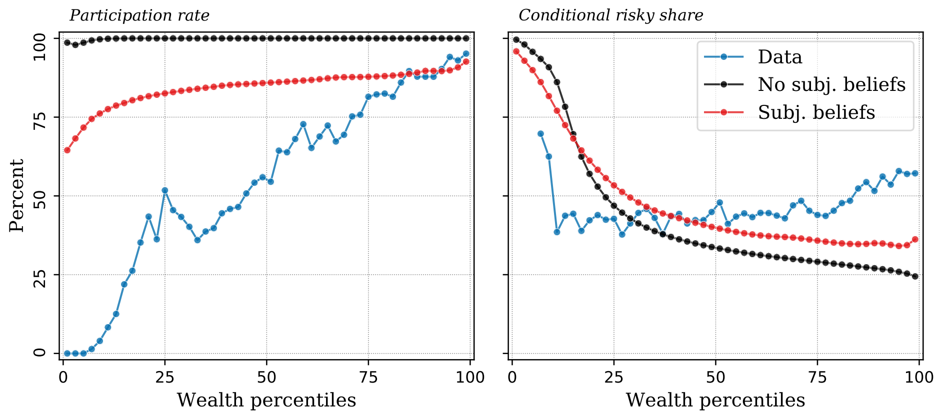


Figure 1.34: Portfolio composition along the wealth distribution. Parameters as reported in Table 1.7.

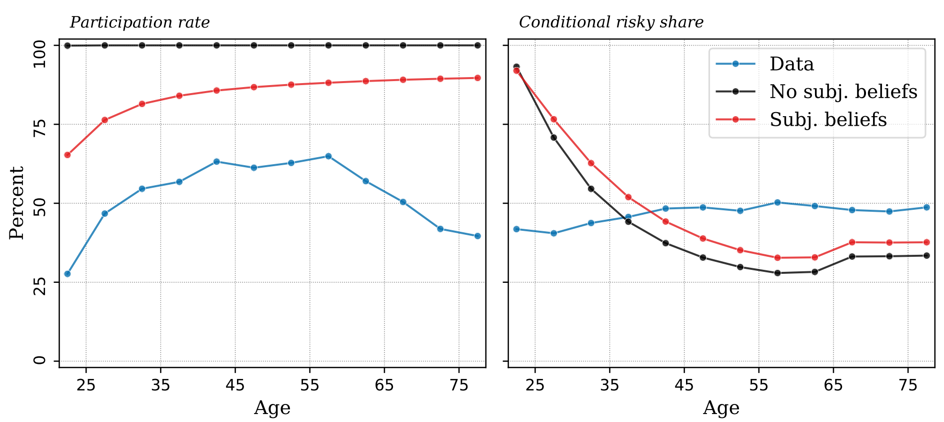


Figure 1.35: Portfolio composition over the life-cycle. Parameters as reported in Table 1.7.

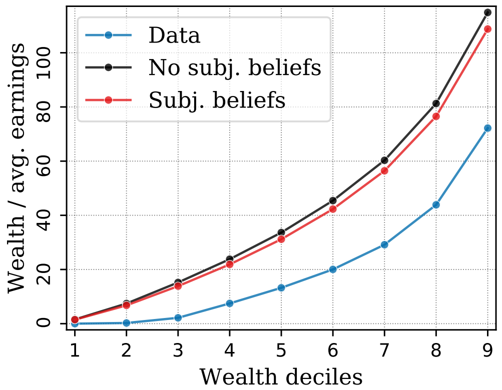


Figure 1.36: Wealth distribution obtained using "standard" calibration. Each dot shows the mean wealth conditional on being in a given decile, normalized by average quarterly earnings. Parameters as reported in Table 1.7. Data source: SCF 1998–2007.

dently, both models miss the empirical wealth distribution by a wide margin. For example, median wealth is approximately three times higher in the model than in the data.

models do even worse when it comes to matching the 10th decile, as there is no mechanism present to generate the thick right tail of the wealth distribution.

1.D Results for the benchmark model with sophisticated households

This section reports the results for the benchmark calibration, but assuming that households are sophisticated and thus anticipate that they will update their beliefs in the future as new return realizations are observed. I focus on the case with per-period participation costs since the differences between the naive and the sophisticated models are very similar in both cases. Figure 1.37 shows the cross-sectional composition generated by the model with sophisticated agents, and how it compares to the benchmark calibration with naive agents. Figure 1.38 repeats the exercise for portfolio allocations over the life-cycle. Figure 1.39 shows that the beliefs do not differ to any considerable extent between naive and sophisticated households which is due to the assumption that investors update their beliefs independently of their participation decision. Figure 1.40 illustrates that while the optimal risky share chosen by naive vs. sophisticated households can differ substantially, these differences predominantly arise for high cash-at-hand levels and therefore have small effects on the portfolio composition in the cross-section.

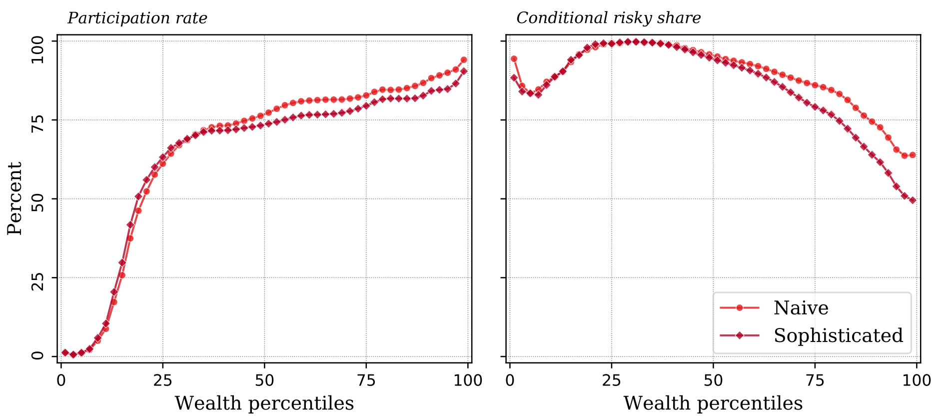


Figure 1.37: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs, naive vs. sophisticated households with subjective beliefs.

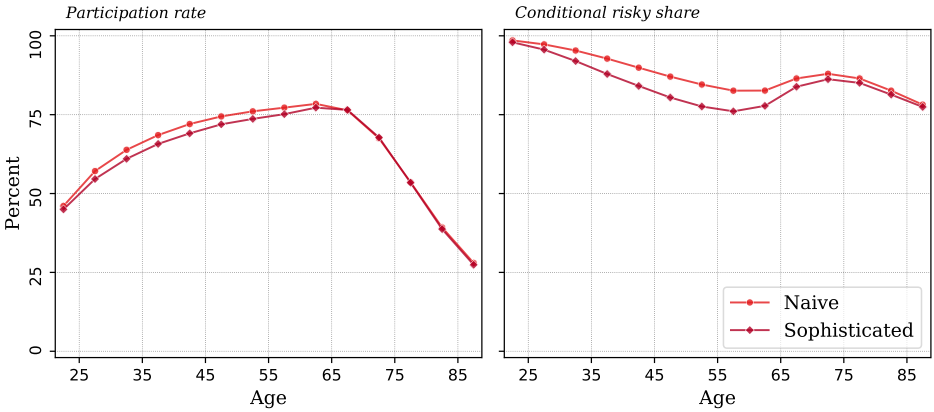


Figure 1.38: Portfolio composition over the life-cycle. Benchmark calibration *with* participation costs, naive vs. sophisticated households with subjective beliefs.

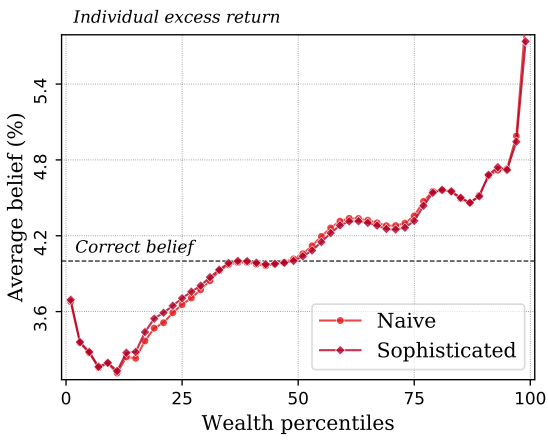
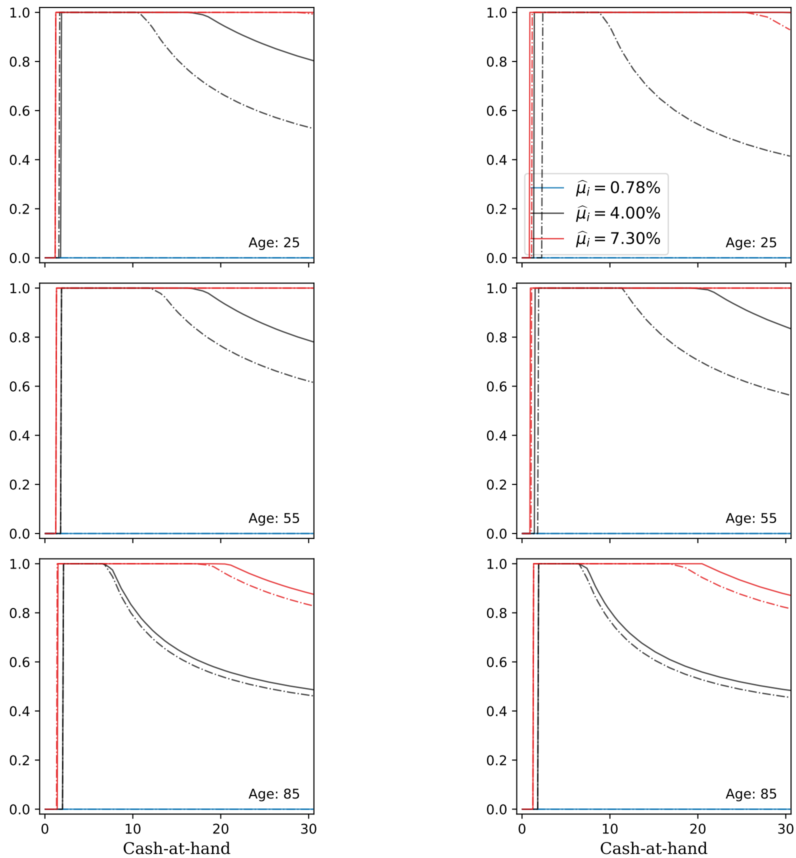


Figure 1.39: Average beliefs about excess returns, naive vs. sophisticated households.



(a) Preference type 1

(b) Preference type 2

Figure 1.40: Optimal risky share for naive (solid line) vs. sophisticated (dashed line) households, shown for a household with the median persistent labor productivity. Impatient households (type 1) in the left-hand column, patient households (type 2) to the right. Colors denote different beliefs about excess returns.

1.E Numerical solution

Household problem

Solving for optimal choices. I solve the household problem using a hybrid endogenous grid point method (EGM). Plain EGM was originally proposed to solve the consumption-savings problem without a portfolio choice, but can be extended to portfolio-choice problems in a straightforward way by adding a root-finding step at each exogenous savings level.

Root-finding needs to be performed on the first-order condition with respect to ξ, the optimal risky share. For the case of EZW preferences, this portfolio Euler equation (P-EE) is given by

$$\pi_h^s \mathbf{E}_i \left[(V')^{-\gamma} V_1' \left(R' - R_f\right) \right] + (1 - \pi_h^s) \mathbf{E}_i \left[((V^b)')^{-\gamma} (V_1^b)' \left(R' - R_f\right) \right] = 0$$

for the working-age household, where the continuation values are given by

$$\begin{aligned} V' &\equiv V_{jh+1}\left(a', p', \widehat{\mu}_i'\right) \\ (V^b)' &\equiv V_j^b\left(a_b'\right) \end{aligned}$$

and their derivatives with respect to to the first argument (i.e., cash-at-hand) are denoted as

$$\begin{aligned} V_1' &\equiv \frac{\partial V_{jh+1}'\left(a', p', \widehat{\mu}_i'\right)}{\partial a'} \\ (V_1^b)' &\equiv \frac{\partial V_j^b\left(a_b'\right)}{\partial a_b'} \end{aligned}$$

Using the envelope condition, the P-EE can be stated as

$$\pi_h^s \mathbf{E}_i \left[(V')^{\psi-\gamma} (c')^{-\psi} \left(R' - R_f\right) \right] + (1 - \pi_h^s) \phi_j^{\frac{1-\gamma}{1-\psi}} \mathbf{E}_i \left[\left(a_b'\right)^{-\gamma} \left(R' - R_f\right) \right] = 0 \quad (1.18)$$

As in plain EGM, I solve for the optimal solution conditioning on an exogenously imposed savings level $b > 0$ (the portfolio choice is indeterminate for $b = 0$, so the boundary case can be ignored). Then (1.18) is an implicit function of ξ via the continuation values, their derivatives and tomorrow's consumption c', which depend on the optimal choice of ξ via its effect on a' and a_b'. It turns out that (1.18) is in general a monotonically decreasing function of ξ, a fact that can be used to easily identify corner solutions as well as interior portfolio choices. Once the optimal portfolio choice has been determined, one can compute the return on the total portfolio $R_{p,t+1}$ which enters the consumption-savings Euler equation. The remaining steps are the same as in the standard EGM procedure.

In the presence of fixed participation costs, EGM has the problem that the first-order conditions are only necessary, but no longer sufficient. Intuitively, this is the case because while the above hybrid method can be used to find an interior optimal solution *conditional on paying the participation cost*, not saving in the risky asset and thus not incurring the participation cost might still yield a higher utility. This potentially creates downward jumps in the consumption policy function which plain EGM cannot deal with. I implement the approach suggested in Iskhakov et al. (2017) and Druedahl and Jørgensen (2017) to address this issue (the variant without taste shocks).

Exogenous grids. I discretize cash-at-hand and the exogenous savings grid in the usual way, putting more points in the region where policy and value functions have more curvature.

I approximate mean beliefs on an age-specific grid of 23 points as follows: I compute the percentiles (P0, P1, P2.5, P7.5, …, P92.5, P97.5, P99, P100) of the cross-sectional belief distribution for each age, and place grid points at the expected values conditional on falling into the regions bracketed by these percentiles. Thus, the grid points are automatically placed in regions where the beliefs for each cohort are concentrated.

The persistent earning component is approximated using the Rouwenhorst procedure with seven grid points. Transitory earning shock realizations are discretized to five possible realizations.

Simulation

Following an approach commonly used in the literature (see, for example, Gomes and Michaelides (2005)), I simulate the distribution of a "representative" cohort over its entire life-cycle instead of rolling forward a cross-section of a limited number of households.

The initial distribution of newborns is computed as follows:

1. Their initial wealth is assigned according to the distribution of wealth in the SCF for the age group 20–25, discretized onto 500 bins.

2. The initial distribution over persistent labor productivity states is the ergodic distribution implied by the Markov process obtained from the Rouwenhorst procedure.

3. Beliefs are drawn from the normal distribution that is assumed for newborns, discretized onto the 23 grid points as outlined above.

Using the households policy rules and the exogenous transition processes, I then build a (sparse) transition matrix that transforms the distribution of households over

exogenous and endogenous states at age t into the corresponding distribution at age $t + 1$. This process is repeated for all ages $t = 0, \ldots, T - 1$. Since there are no endogenous intergenerational linkages (such as bequests between a specific parent and descendant household), the stationary distribution over all cohorts can then be obtained as a cohort-size-weighted average of these age-specific distributions.

Chapter 2

Health Dynamics and Heterogeneous Life Expectancies

(joint with Jonna Olsson)

2.1 Introduction

In this paper, we provide improved estimates for age-dependent health transitions and survival probabilities for different subsamples of the U.S. population. The estimated yearly transition matrices for health and death can be used in any life-cycle model where the evolution of health and survival probability is of interest. The implied life expectancy, which does not only depend on the standard dimensions of gender, race, and age, but also on health, can be used to assess inequality in longevity in the population based on health and potentially also other characteristics of the individual.

Numerous studies have identified health dynamics and health shocks as a major source of risk over the life cycle. A negative health shock can result in large medical expenditures (De Nardi, French, and Jones 2010; Kopecky and Koreshkova 2014), which affects the incentives to accumulate assets, and could also affect the earnings potential (French 2005; Coile, Milligan, and Wise 2016). The survival probability directly affects the effective discount factor, a mechanism present in any life-cycle model with an uncertain life span (see De Nardi, Pashchenko, and Porapakkarm 2017 for estimates of this effect as well as a comprehensive estimate of the cumulative effects of bad health). According to Finkelstein, Luttmer, and Notowidigdo 2013, the health state directly influences the marginal utility from consumption. Hence, in order to quantify the risk an individual faces and model the choices and actions the individual takes, a realistic health and survival process is crucial.

We use the Health and Retirement Study (HRS), a representative panel of elderly U.S. households, to investigate the development of health and longevity in the later stages of life. The survey includes, among other things, information on self-reported health and the date of death, if applicable. The survey started in 1992, and many of the respondents have died over the sample period, making it the preferred data set for studying survival and health dynamics.

However, the data set also has some peculiarities. For most cohorts and time periods, the survey has been conducted on a biennial basis. In practice, due to variation in interview dates, we effectively observe individual time spans of one, two or three years for each interview wave.[1] Second, the death dates are coded exactly, and do not follow the interview wave structure. Third, even if all observations were perfectly biennial, the observations are overlapping: we observe transitions for one person at ages a and $a + 2$ and for another person at ages $a + 1$ and $a + 3$. Both observations should contribute to estimating the one-year health and survival transitions between ages $a + 1$ and $a + 2$.

In this paper, we estimate a yearly five-state Markov chain that can be used in a life-cycle model. Our methodology is similar in spirit to that of Pijoan-Mas and Ríos-Rull 2014, but takes into account the irregular and overlapping observations described above and estimates *one-year* health transition probabilities and survival probabilities that depend on sex, race, age and health status (as opposed to previous studies using the HRS, which estimate two-year probabilities).

Conceptually, our method is a straightforward maximum likelihood estimator, where we maximize the probability of observing the *transition paths* in data. Each transition path consists of an initial health state, a number of periods, and an end point, which is either a health state (conditional on survival) or death. The estimation includes rolling forward each start observation the appropriate number of years with the Markov chain that is being estimated. To put structure on the Markov chain we use a nested logit, where survival and health transitions conditional on survival are modeled as functions of the current health state and age. The probability of survival follows the usual binary-outcome logit model, while, conditional on survival, health transitions are modeled using multinomial logit.[2]

Previous estimates of health transitions and survival probabilities in the literature have been restricted by data limitations. In terms of coverage, the HRS is arguably the richest data set, but as explained, the data structure is biennial at best. The estimates based on the HRS, e.g., Pijoan-Mas and Ríos-Rull 2014 or Amengual, Bueren, and Crego 2017, have therefore been two-year transition probabilities.[3] However, most life-cycle applications need an annual Markov chain, and consequently, some authors have resorted to the Panel Study of Income Dynamics (PSID), which up until 1997 was annual (e.g., French 2005 or De Nardi, Pashchenko, and Porapakkarm 2017). This panel data set covers a long time period, but the number of individuals is relatively

1. Moreover, even though attrition is low, there are individuals who are not observed in one or more waves, so the time span between two observations can exceed three years.

2. We also implemented an alternative specification, estimating the health and survival probabilities separately in a two-step procedure, which yields very similar results.

3. Some studies have calibrated the health dynamics in an annual model to HRS data, but then used much simpler specifications, e.g., French and Jones 2011 use a two-state health process where the health status and mortality depend on previous health status interacted with an age polynomial.

small. Due to too few observations it is not possible to keep the full level of detail using all five health states and it has been shown that the PSID data set severely underestimates mortality rates.[4] Moreover, the PSID is also biennial from 1997 and onward, so more recent developments in mortality and health cannot be included in an estimation based on annual data.

The method we propose overcomes these challenges that previous literature has struggled with and estimates annual transition probabilities, based on the rich HRS data and keeping the full state space with five health states.

The resulting health gradient for longevity is strong. For a 70-year-old nonblack man in excellent health, the probability of reaching his 80th birthday is around 75%, while the corresponding probability for a nonblack man in poor health is just below 40%. Another way of expressing the same health gradient is that the expected longevity for a 50-year-old nonblack man in excellent health is 79 years, while it is only 73 years for a man in poor health.

The relationship between socio-economic status and life expectancy is well established but remains poorly understood (see, e.g., Chetty et al. 2016 and the references therein). With our methodology we can include time-invariant characteristics when estimating the health and survival dynamics. We show that there is substantial inequality in life expectancy between different educational groups. The average nonblack man with less than a high school degree has a life expectancy of 75 years at the age of 50, while the average for nonblack men with some college education or more is 80 years. This difference is due to two factors. First, at the age of 50, overall health is worse in the group with lower education. Second, even conditional on health status, the health dynamics and survival probabilities for this group are worse also from the age of 50 and onwards. We estimate that of the differences in life expectancy between the low and high educated, approximately one fifth is due to worse overall health at the age of 50, while the lion's share is due to worse health and survival dynamics after that age.

Most estimates of health processes used for life-cycle models collapse the state space for health into two groups: good or bad health (French 2005; French and Jones 2011; De Nardi, Pashchenko, and Porapakkarm 2017).[5] There are two benefits from using the full state space as reported in the HRS: First, trivially, a larger state space captures more of the heterogeneity in the population. Second, a richer state space allows for more correct dynamics and persistence of the process. The drawback of using more states is, of course, the additional computational burden. However, we think that it is key to properly capture the persistence and duration dependence of

4. French 2005 combines the PSID data with mortality statistics from the National Center for Health Statistics, since the PSID data underestimates mortality rates by 25%.

5. One reason is that to estimate yearly transitions, authors have resorted to using PSID, which until 1997 was a yearly survey. However, the number of individuals there is relatively small and therefore it is necessary to combine data into coarser health states.

staying in bad health when modeling the health risk that individuals face (see De Nardi, Pashchenko, and Porapakkarm (2017) for a discussion).

Our estimated process is highly persistent, especially for the worst health state. Once there, the probability of remaining in the worst health state another period is above 75%. The importance of health persistence is stressed by, e.g., Contoyannis, Jones, and Rice 2004, and the persistence we estimate is well in line with the persistence for the five-state frailty index that Hosseini, Zhao, Kopecky, et al. 2018 develop.[6]

If we use our five-state process, but interpret the results according to a two-state classification (with the two worst health states classified as "bad", in line with previous literature), the result is a negative duration dependence in the probability of recovering from bad health.[7] For a 70-year-old nonblack man who has been in bad health only one year, the probability of recovering to good health is 20%, while if he has been in bad health five years, the probability is five percentage points lower.

In the next section we describe the structure of the HRS data, and thereafter we describe the estimation in detail. Section four shows the baseline results while section five analyzes the health and life expectancy inequality between different education groups. The last section concludes.

2.2 Data

We use the Health and Retirement Study (HRS), a representative panel of U.S. households in older ages, to investigate the development of health and longevity in the later stages of life. The survey includes, among other things, questions about self-reported health state and date of death, if applicable.

The survey started in 1992 and in this paper, we use HRS data up to and including the eleventh wave in 2012.[8] The first cohort included in the survey was between 51 and 61 years old in 1992, and thereafter new cohorts have been added. Many of the respondents have died over the sample period, making it an appropriate data set for

6. It is also substantially higher than the persistence which Hosseini, Zhao, Kopecky, et al. 2018 calculate for self-reported health based on PSID. This points both at the strength of HRS (which includes more individuals in older ages), and at the importance of estimating age-dependent transition probabilities.

7. Another alternative is to, like De Nardi, Pashchenko, and Porapakkarm 2017, use a two-state health process with a second-order Markov process and fixed ex-ante heterogeneity to capture the duration dependence.

8. The RAND version O, covering waves up until 2012, is the most recent RAND release that includes data from the National Death Index (NDI). Since correct death dates are crucial for our analysis, this is our preferred data set. There is one later RAND release, covering the 2014 wave as well, but there NDI data is lacking. An analysis shows that there are discrepancies in death dates between the exit interview information and the NDI date of death.

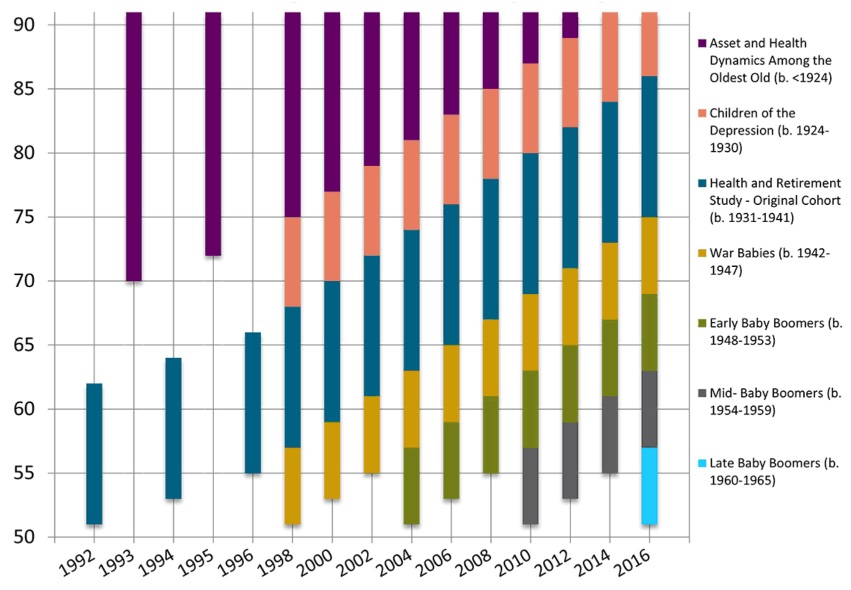


Figure 2.1: Structure of the Health and Retirement Study. The y-axis shows the age of the respondents in each cohort and wave. Graph taken from https://hrs.isr.umich.edu/documentation/survey-design, where additional information about the survey design is available.

studying survival. Figure 2.1 shows the panel structure of the HRS and how new cohorts have been incorporated into the survey over time.

As can be seen from Figure 2.1, the survey has been conducted biennially for most cohorts and time periods. However, in practice, there is a substantial variation in the time elapsed between interviews. Each survey round is conducted over a period of time, and the actual time elapsed between interviews for a respondent for two consecutive waves varies between one and three years. For respondents missing one or more interviews, the time interval between two interviews or the time elapsed between the last interview and the death date is more than three years. Figure 2.2 shows the distribution of the actual time elapsed from one observation to the next for the full sample. As can be seen, slightly more than 80% of the transitions are best characterized as two-year transitions, but almost 20% are not. Figure 2.3 illustrates two typical observations in our sample.

The two key variables we use are self-reported health and date of death. Self-reported health is simply the respondent's answer to the question "Would you say your health is excellent, very good, good, fair, or poor?". The first reason to use this variable is the general availability of this information. Very similar questions

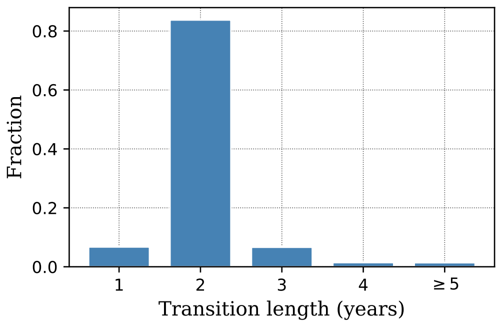


Figure 2.2: Time elapsed between two observations (second observation being either a new interview or the death date).

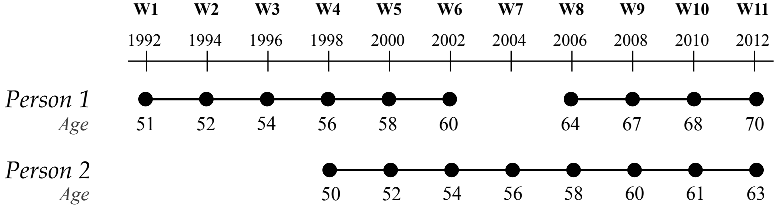


Figure 2.3: Two individuals in the HRS, illustrating the irregular and missing observations. Waves and calender years indicate the biennial structure, age indicates the actual age at the time of the interview.

are asked in many other surveys, both within the U.S. (for example the Panel Study of Income Dynamics (PSID) and the Medical Expenditure Panel Survey (MEPS) include this question) and also globally (for instance the Survey of Health, Aging and Retirement in Europe (SHARE) asks about self-reported health). Hence, the insights into the dynamics of self-reported health and life expectancy conditional on this measure can be used for analyses based on many other data sets.

Second, a number of studies have shown that self-reported health is highly correlated with other subjective and objective measures of health and is also a good predictor for future mortality (see e.g. Idler and Benyamini 1997 and Pijoan-Mas and Ríos-Rull 2014). Self-reported health can be interpreted as a one-dimensional variable capturing high-dimensional information, letting the respondent aggregate this information him- or herself.[9]

Figure 2.4 shows the distribution of nonblack individuals by health state for different ages.[10] The overall health is declining in age, but it might be surprising that the health distribution among 50-year-old individuals is not *that* much better than among 90-year-old individuals. This suggests that the aggregation of underlying health measures done by the respondent also takes into account the relative health within cohort. A 70-year-old respondent who reports "excellent" health can still feel worse than his/her 20-year self, but "excellent" in comparison to what the person perceives could be expected as a 70-year-old. Since all our estimates will be conditional on age this is taken into account.

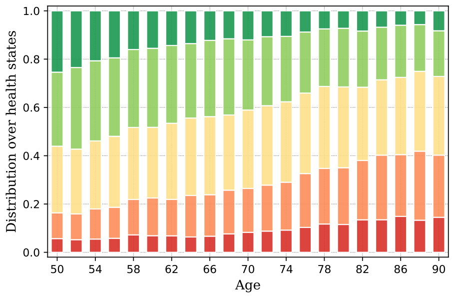
The second key variable is the date of death. An important feature of the HRS data is that the death date is recorded, regardless of whether the respondent stayed in the survey until his/her death or not. Hence, the attrition for this particular information is virtually zero. The death date is first recorded in the so-called exit interview or by the surviving spouse. The survey is then complemented with information from the National Death Index (NDI). Therefore, a correct date of death is recorded even for those who dropped out of the survey, and the death date does not follow the biennial wave structure but corresponds to the actual death date.

We exclude all observations with missing age, race, gender or self-reported health, and those where we only have one observation for the individual (since then we do not have any transition probabilities to estimate). Further, we restrict the sample to individuals above the age of 50.[11] We estimate the subsamples of men/women and the nonblack/black population separately, since it is well known that the life expectancies

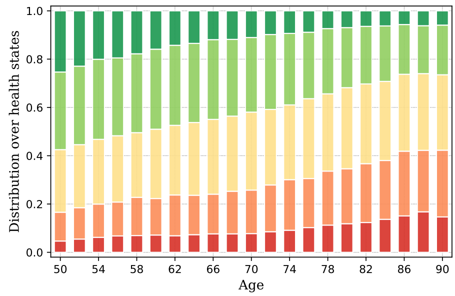
9. An alternative is to let the researcher do the aggregation into a single index, incorporating different physical and mental conditions. See Amengual, Bueren, and Crego 2017 and Hosseini, Zhao, Kopecky, et al. 2018 for other suggested health groupings.

10. Black individuals are generally slightly worse in terms of health. Corresponding graphs for black men and women can be found in the appendix.

11. Even though no one in the core sample should be under that age, the full sample includes spouses that could be younger.



(a) Nonblack men



(b) Nonblack women

Figure 2.4: Distribution of health states by age. Red color indicates worst ("poor") health state while dark green indicates best ("excellent") health state. Observations are grouped into two-year age bins.

		N. of indiv.	N. obs.	Avg. obs./indiv.	Age		
					Min.	Mean	Max.
Male	Non-black	12,765	77,612	6.1	50	67.7	106
	Black	2,397	12,252	5.1	50	65.9	103
Female	Non-black	15,485	99,455	6.4	50	68.4	110
	Black	3,556	19,556	5.5	50	66.4	111

Table 2.1: Sample description.

for these subgroups follow very different trajectories, and are reported separately by the NVSS.[12] Table 2.1 shows the number of individuals and observations by subgroup. We show estimates for nonblack men and women in the main document, due to the larger sample size. The results for black men and women are available upon request. For all estimations we weight observations by their person-level analysis weight.

12. The HRS also has information about Hispanic origin, but the Hispanic group was not added to the annual life table program until 2006, while we use earlier data for some comparisons (see Arias 2014 for technical notes describing the life table program). Therefore, we follow the older life table classifications and only use the black/nonblack categories for our estimates (with Hispanics included in the nonblack group). Note also that other groups than white and black (and later Hispanics) are not reported separately by the NVSS, due to concerns about data limitation and misclassification. We keep the label "nonblack" throughout the document, but when comparing results to life tables we make the comparison with the "white" group.

2.3 Estimation

Our goal is to estimate a Markov chain for *annual* survival probabilities and health-to-health transitions conditional on survival. The approach takes into account that observations in the HRS can occur at irregular (mostly biennial but also non-biennial) frequencies.[13] Our proposed methodology allows us to address two shortcomings of the HRS data:

(1) We need to estimate annual transition probabilities, but HRS waves are biennial. Due to variation in interview dates, we effectively observe transitions over one, two, three or more years, with about 80% of transitions being best described as two-year transitions.

(2) Additionally, transition intervals are overlapping: we often observe transitions between ages a and $a + 2$ for one respondent, while observing a transition from $a + 1$ to $a + 3$ for another respondent. These two observations contribute to estimating the transition matrices between ages $a + 1$ and $a + 2$. Therefore, we cannot simply estimate two-year transition matrices and take a matrix "square root" to obtain one-year transitions.

We implement a custom maximum-likelihood estimator that takes into account varying transition frequencies and overlapping transition intervals. In this section we describe an approach that jointly estimates the parameters governing both health-to-health and survival transitions in a single step.[14]

Transition probabilities

While the HRS itself is organized into individual/year observations, for the purpose of estimating transition probabilities we will reinterpret the sample such that one *transition* constitutes one observation. Each transition consists of an observed starting health state, a number of periods, and an end point, which is the next time the individual is observed. The end point can be either a new health state or death. In what follows, we index transitions by i but usually skip writing the index explicitly.

We refer to a transition's starting date as $\underline{t}$, to its length (in years) as T, and to its end date as $\overline{t} = \underline{t} + T$. We denote by the tuple $(h_t, \boldsymbol{x}_t)$ the information available at time t, $t \in \{\underline{t}, \dots, \underline{t} + T\}$, where $h_t \in \{1, \dots, H\}$ is an individual's self-reported health state, with 1 representing the best and H the worst realization. The vector $\boldsymbol{x}_t$

13. This is in contrast to Pijoan-Mas and Ríos-Rull (2014), who estimate two-year transition probabilities and force all transitions to be at a biennial frequency regardless of the actual time passed between two consecutive interviews, and furthermore restrict transitions to occur between even ages, i.e., they assume that the observed transitions line up with ages $50, 52, 54, \dots$.

14. As a robustness test, we also estimate the process in two separate steps: first health-to-health transitions conditional on survival, and then survival probabilities conditional on health. The results are very similar and available upon request. Since the nested version uses more information, this is our preferred method.

contains any other variables of interest, in particular age. We allow for time-invariant characteristics such as birth year, gender, race or education level to be included in $\boldsymbol{x}_t$, but restrict the time-varying variables to age and potentially calendar year. This restriction is necessary as we need to compute the evolution of $\boldsymbol{x}_t$ over $\underline{t}+1, \underline{t}+2, \ldots$ for multi-year transitions, which is not possible in general except for variables that follow a deterministic path (such as age and calendar year).

Let s_t be a binary indicator for whether a person is alive at date t,

$$s_t = \begin{cases} 1 & \text{if alive at } t \\ 0 & \text{else} \end{cases} \tag{2.1}$$

We assume that the one-period-ahead probability of survival is given by the binary-outcome logit model

$$p^s_{t+1} \equiv \Pr\left(s_{t+1} = 1 \mid h_t, \boldsymbol{x}_t \right) = \frac{1}{1 + e^{-g(h_t, \boldsymbol{x}_t \mid \gamma)}} \tag{2.2}$$

Survival probabilities are governed by the parameter vector γ which is to be estimated. Similarly, conditional on survival, the probability that health state j is realized next period is given by the multinomial logit formula

$$p^{h,j}_{t+1|s} \equiv \Pr\left(h_{t+1} = j \,\middle|\, s_{t+1} = 1, h_t, \boldsymbol{x}_t \right) = \frac{e^{f_j(h_t, \boldsymbol{x}_t \mid \boldsymbol{\beta}_j)}}{\sum_m e^{f_m(h_t, \boldsymbol{x}_t \mid \boldsymbol{\beta}_m)}} \tag{2.3}$$

with parameter vector $\boldsymbol{\beta}_j$ to be estimated for each outcome j.[15] Here we use the notation $p^h_{\bullet|s}$ to indicate that this probability is conditional on survival. We can then compute the unconditional probability of being in health state j in the next period as

$$p^{h,j}_{t+1} = p^{h,j}_{t+1|s} \times p^s_{t+1} \tag{2.4}$$

Below we will frequently want to emphasize that we condition on a particular health state $h_t = k$, and hence we will use the expressions

$$p^s_{t+1|k} \equiv \Pr\left(s_{t+1} = 1 \mid h_t = k, \boldsymbol{x}_t \right) \tag{2.5}$$

$$p^{h,j}_{t+1|k,s} \equiv \Pr\left(h_{t+1} = j \,\middle|\, s_{t+1} = 1, h_t = k, \boldsymbol{x}_t \right) \tag{2.6}$$

which are otherwise identical to those established earlier.

In the following sections we lay out the estimation strategy to determine the parameter vectors γ and $\boldsymbol{\beta}_j$ for all outcomes j. We defer stating the exact functional forms of $g(\bullet)$ and $f(\bullet)$, as these are specific to the particular model to be estimated (e.g. whether cohort fixed effects or education are included, etc.).

15. We assume that all parameters in $\boldsymbol{\beta}_j$ are specific to outcome j and there are no "common" parameters shared across all outcomes. This is due to the fact that we have no outcome-specific regressors and thus any common parameters would cancel out in (2.3), leaving these parameters unidentified.

Maximum-likelihood approach

In this section we describe the ML approach that jointly estimates the parameters governing both health-to-health and survival transitions in a single step.

An illustrative example. Before deriving the probabilities that should be plugged into the log-likelihood function, it is worthwhile to work through an illustrative example for the case of the nested logit estimator. We consider a simplified setup with only two health states and assume a two-year transition, as illustrated in Figure 2.5.

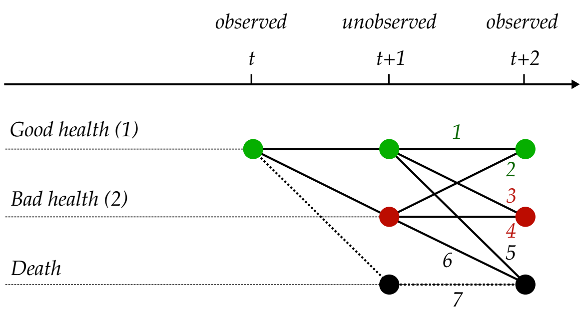


Figure 2.5: Simplified transition paths

At $t+2$ there are three possible outcomes, but seven distinct paths via which these outcomes can be realized. It turns out that the probability distribution that should enter the likelihood function is one over paths, not outcomes. To make this point, first consider a PMF over outcomes in $t+2$, which is given by the three probabilities

$$\Pr\left(h_{t+2} = 1 \,|\, h_t\right)$$
$$\Pr\left(h_{t+2} = 2 \,|\, h_t\right)$$
$$\Pr\left(s_{t+2} = 0 \,|\, h_t\right)$$

For either health state $j = 1, 2$ these can be computed as follows:

$$\Pr\left(h_{t+2} = j \,|\, h_t\right) = \sum_{m=1}^{2} \Pr\left(h_{t+2} = j \,|\, h_{t+1} = m, h_t\right) \Pr\left(h_{t+1} = m \,|\, h_t\right)$$

On the other hand, the probability of observing death in $t+2$ can be written as

$$\Pr\left(s_{t+2} = 0 \,|\, h_t\right) = \sum_{m}^{2} \Pr\left(s_{t+2} \,|\, h_{t+1} = m, h_t\right) \Pr\left(h_{t+1} = m \,|\, h_t\right) + \Pr\left(s_{t+1} = 0 \,|\, h_t\right)$$

The issue with this formulation is that whenever death in $t+2$ is observed, the probability of this outcome includes the case that the individual already died in $t+1$,

which corresponds to path *7* in Figure 2.5. However, due to how the transition data was constructed this is impossible, as a one-period observation would have been recorded if an individual had already died in $t+1$ (remember that the date of death is always recorded correctly and does not follow the wave structure). Hence, the probability associated with path *7* should never enter the likelihood function. This issue becomes even more pronounced for longer transitions, since the probability of ending up in the absorbing death state in the penultimate period is strictly increasing in the transition length.

To properly address this issue, we instead propose to compute the distribution over paths instead of outcomes. Naturally, in the above example we do not know whether path *1* or *2* was realized when we observe the outcome $h_{t+2} = 1$, so the probabilities of both will have to be included in that case, and analogously for the remaining outcomes.

To shut down all paths leading to "premature" death before the terminal period (which is only path *7* in the above example), we want to evaluate the probabilities of the events

$$\Pr\left(h_{t+2} = 1 \wedge s_{t+1} = 1 \,|\, h_t \right)$$
$$\Pr\left(h_{t+2} = 2 \wedge s_{t+1} = 1 \,|\, h_t \right)$$
$$\Pr\left(s_{t+2} = 0 \wedge s_{t+1} = 1 \,|\, h_t \right)$$

For either health outcome $j = 1, 2$ we find that

$$\begin{aligned} \Pr\left(h_{t+2} = j \wedge s_{t+1} = 1 \,|\, h_t \right) &= \Pr\left(s_{t+1} = 1 \,|\, h_{t+2} = j, h_t \right) \times \Pr\left(h_{t+2} \,|\, h_t \right) \\ &= \Pr\left(h_{t+2} \,|\, h_t \right) \end{aligned}$$

which follows since

$$\Pr\left(s_{t+1} = 1 \,|\, h_{t+2} = j, h_t \right) = 1$$

An individual who is in health state j at $t+2$ must have been alive at $t+1$, so the additional restriction that $s_{t+1} = 1$ is redundant for health outcomes. However, this is not the case for the probability of being dead in $t+2$:

$$\begin{aligned} \Pr\left(s_{t+2} = 0 \wedge s_{t+1} = 1 \,|\, h_t \right) &= \sum_{m=1}^{2} \Pr\left(s_{t+2} = 0 \wedge s_{t+1} = 1 \,|\, h_{t+1} = m, h_t \right) \\ &\qquad \times \Pr\left(h_{t+1} = m \,|\, h_t \right) \\ &= \sum_{m=1}^{2} \Pr\left(s_{t+2} = 0 \,|\, h_{t+1} = m, h_t \right) \Pr\left(h_{t+1} = m \,|\, h_t \right) \end{aligned}$$

The second line follows since conditional on $h_{t+1} = m$, we necessarily have $s_{t+1} = 1$. This formulation shuts down any paths with $s_{t+1} = 0$. Note that we can still compute

the probability of such paths and include them in the log-likelihood formula, but since these are never observed, they will never be selected to actually contribute to the log-likelihood function. The consequence is that we can entirely skip computing the probability of these paths, which also leads to the at first puzzling corollary that for the components that actually do enter the log-likelihood, we have

$$\Pr\left(h_{t+2} = 1 \,|\, h_t \right) + \Pr\left(h_{t+2} = 2 \,|\, h_t \right) + \Pr\left(s_{t+2} = 0 \wedge s_{t+1} = 1 \,|\, h_t \right) < 1$$

in general.

General setup. We now return to the issue of estimating the parameter vector $\boldsymbol{\theta}$ that governs health-to-health and survival transitions, defined as

$$\boldsymbol{\theta} \equiv (\boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_m, \ldots, \boldsymbol{\beta}_H, \boldsymbol{\gamma}) \in \mathbb{R}^K$$

where $K = (H-1)K_h + K_s$, $\boldsymbol{\beta}_m \in \mathbb{R}^{K_h}$ for each m as before and $\boldsymbol{\gamma} \in \mathbb{R}^{K_s}$. We omit the normalized base outcome parameter vector $\boldsymbol{\beta}_1 = \mathbf{0}$ for health state 1. From any transition bracketed by the dates $\underline{t}$ and $\overline{t}$ we obtain one observation, a PMF over health states "augmented" by the state of death. We call this vector $\boldsymbol{\mu}_t \in \mathbb{R}^{H+1}$. In $\underline{t}$ we impose the degenerate initial distribution

$$\boldsymbol{\mu}_{\underline{t}} = (\underbrace{0, \ldots, 0, 1, 0, \ldots, 0}_{H \text{ elements}}, 0)^\top \tag{2.7}$$

with unity in the position corresponding to the initial health state $h_{\underline{t}}$.

The one-year health-to-health transition matrix conditional on survival is given by

$${}^{h}_{t}(\boldsymbol{x}_t|\boldsymbol{\beta}) = \begin{bmatrix} p^{h,1}_{t+1|1,s} & \cdots & p^{h,H}_{t+1|1,s} \\ \vdots & \ddots & \vdots \\ p^{h,1}_{t+1|H,s} & \cdots & p^{h,H}_{t+1|H,s} \end{bmatrix} \tag{2.8}$$

where the conditional probabilities $p^{h,j}_{t+1|k,s}$ are defined in the same way as in (2.6). This transition matrix is a function of the covariate vector $\boldsymbol{x}_t$ and possibly calendar time t, but not of the current health state h_t as it contains transitions for all possible current h_t.

Let $\boldsymbol{\pi}^s_t$ be the vector of survival probabilities between periods t and $t+1$ for each health state $k \in \{1, \ldots, H\}$ today,

$$\boldsymbol{\pi}^s_t\left(\boldsymbol{x}_t|\boldsymbol{\gamma}\right) = \left(p^s_{t+1|1}, \ldots, p^s_{t+1|k}, \ldots, p^s_{t+1|H}\right)^\top \tag{2.9}$$

where any element $p^s_{t+1|k}$ is obtained as stated in (2.2). Given the distribution over health states *conditional* on being alive in t, $\boldsymbol{\mu}^h_t$, the probability of being alive in $t+1$

is therefore

$$p^s_{t+1}(\gamma) = \boldsymbol{\pi}^s_t(\gamma)^\top \boldsymbol{\mu}^h_t \tag{2.10}$$

We can now write down the joint health/survival transition matrix, given by

$$_t(\boldsymbol{x}_t|\boldsymbol{\theta}) = \begin{bmatrix} p^{h,1}_{t+1|1,s}p^s_{t+1|1} & \cdots & p^{h,H}_{t+1|1,s}p^s_{t+1|1} & \left(1 - p^s_{t+1|1}\right) \\ \vdots & \ddots & \vdots & \vdots \\ p^{h,1}_{t+1|H,s}p^s_{t+1|H} & \cdots & p^{h,H}_{t+1|H,s}p^s_{t+1|H} & \left(1 - p^s_{t+1|H}\right) \\ 0 & \cdots & 0 & 1 \end{bmatrix}$$

We can then generate the distribution $\boldsymbol{\mu}_t$ over health/death states for any t by repeatedly applying the transition matrix, starting with the degenerate initial distribution (2.7). The law of motion for $\boldsymbol{\mu}_t$ is therefore

$$\boldsymbol{\mu}_{t+1}(\boldsymbol{\theta})^\top = \boldsymbol{\mu}_t(\boldsymbol{\theta})^\top {}_t(\boldsymbol{x}_t|\boldsymbol{\theta}) \tag{2.11}$$

In line with the initial discussion on computing PMFs over outcomes versus realizations of complete paths, we need to discard any paths that pass through the state $s_{\bar{t}-1} = 0$. This can be achieved by computing the PMF $\boldsymbol{\mu}_{\bar{t}-1}$ according to (2.11) and then defining the "pseudo" PMF

$$\widetilde{\boldsymbol{\mu}}_{\bar{t}-1} = \left(\mu_{1,\bar{t}-1}, \ldots, \mu_{H,\bar{t}-1}, 0\right)$$

Note that $\sum_j \widetilde{\mu}_{j,\bar{t}-1} = \Pr\left(s_{\bar{t}-1} = 1 \mid h_{\underline{t}}, \boldsymbol{x}_{\underline{t}}\right)$, the probability of being alive in $T-1$. The terminal distribution of interest can then be computed as before, i.e.,

$$\boldsymbol{\mu}_{\bar{t}}(\boldsymbol{\theta}) = \widetilde{\boldsymbol{\mu}}_{\bar{t}-1}(\boldsymbol{\theta})^\top {}_{\bar{t}-1}(\boldsymbol{x}_{\bar{t}-1}|\boldsymbol{\theta}) \tag{2.12}$$

Log-likelihood. We are now ready to write down the likelihood function for observation i. Let $\delta^{h,j}_t$ be the indicator variable defined as

$$\delta^{h,j}_t = \begin{cases} 1 & \text{if } h_t = j \\ 0 & \text{else} \end{cases} \tag{2.13}$$

and $s_{\bar{t}}$ be the indicator for being alive in $\bar{t}$, analogous to (2.1). Then the likelihood function for transition i is given by

$$\mathcal{L}_i(\boldsymbol{\theta}) = s_{\bar{t}}\left(\sum_{j=1}^{H} \delta^{h,j}_{\bar{t}} \log \mu_{j,\bar{t}}(\boldsymbol{\theta})\right) + (1 - s_{\bar{t}}) \log \mu_{H+1,\bar{t}}(\boldsymbol{\theta}) \tag{2.14}$$

The estimated parameter vector $\widehat{\boldsymbol{\theta}}$ is hence the vector that maximizes the weighted sum of the log-likelihoods over all observations.[16]

16. Even though it is conceptually a standard log-likelihood estimation, the implementation is non-standard and not included in any existing software, but specifically implemented for the problem at hand.

Baseline model specification

In the baseline specification of the functions governing the survival and health transitions, we include age polynomials of degree two as the only covariates other than the current health state, i.e., $\boldsymbol{x}_t = (a_t)$. We thus have the following functional forms for $g(\bullet)$ and $f_j(\bullet)$:

$$g(h_t, \boldsymbol{x}_t|\gamma) = \sum_{n=0}^{2}\sum_{k=1}^{H} \gamma_{nk} \times \delta_t^{h,k} \times a_t^n$$

$$f_j\left(h_t, \boldsymbol{x}_t|\boldsymbol{\beta}_j\right) = \sum_{n=0}^{2}\sum_{k=1}^{H} \beta_{jnk} \times \delta_t^{h,k} \times a_t^n \qquad \forall\ j \in \{2, \ldots, H\}$$

Note that one benefit of using the nested logit formulation is that it is possible to include different covariates in the functions governing the survival process and the health process. However, we find that a second-order polynomial is sufficient in both.

2.4 Baseline results

In this section, we first show the results from the estimations and thereafter, to verify our estimates, we compare the model-predicted transition probabilities to what we observe in the data. We then compute model-predicted life expectancy conditional on health, taking into account potential future health transitions to evaluate the health gradient for survival. Once more, we verify our results, this time by making a comparison to national statistics from NVSS. We conclude with a short discussion on the duration dependence of being in "bad" health, classifying the worst two health states as "bad" (in line with previous literature). We show that our five-state health process captures some of the duration dependence of bad health.

The resulting transition matrices by age, which can be directly incorporated in a life-cycle model, are downloadable from the authors' websites.

Health transitions and survival probabilities

The health-to-health transition probabilities for nonblack males are shown in Figure 2.6, while those for nonblack women are shown in Figure 2.7. As can be seen, the health state is persistent: for a 70-year-old nonblack man in poor health, the probability of remaining in the same poor health state next year is around 75%. The probability of improving to anything better than the second worst health state is low, below 5%.

The same pattern holds true for all health states: to remain in the current health state is the most likely outcome, and to improve or deteriorate one step is the second

most likely outcome. For a 50-year-old nonblack man in the best health state, the probability of remaining in excellent health is around 70%, but as age increases, it becomes more likely to transition to the second best health state.

Figure 2.8 shows the survival probabilities for nonblack men and women conditional on health. Unsurprisingly, the survival probability is decreasing in age, but there is also a clear health gradient. The probability of surviving one year ahead for a 50-year-old in the best health state (excellent) is almost 100% while for an otherwise similar individual in the worst health state it is approximately 95%. As is well known, the survival probability conditional on age is higher for women than for men.

Another way of illustrating the dynamics of the estimated health and survival process is given in Figure 2.9, which shows the evaluation of probabilities for each health state and for being dead by year, given an initial health state and a starting age. As can be seen, the survival probability differs substantially depending on the initial health state: for a nonblack man aged 70 in excellent health, the predicted probability of surviving an additional 10 years is more than 80%, but if he is in poor health, the probability is just around 40%.

Comparing model predictions and data

In order to compare model predictions to "raw" data moments, we compute the two-year transition probabilities implied by our annual model. Then, we compare these to the fraction of individuals with a particular outcome in a subsample restricted to two-year transitions, which is the large majority of observations, as shown in Figure 2.2.

Figure 2.10 shows the model-predicted two-year health-to-health transitions for nonblack men compared to the actual observed two-year transitions, while Figure 2.11 shows the corresponding information for nonblack women. Given the strict functional form assumptions that we impose, the estimated probabilities and the data are remarkably close.

To compare model predictions for survival with observed data, we compare the estimated cumulative probability of having died within two years to the actual observed death within two years since the last observations. Figure 2.12 shows the results. Again, the model predictions and the data are remarkably close.

To assess how well our model predicts long-run outcomes, we compare actual survival rates as observed in the HRS with model predictions over a time horizon of up to 20 years. Figure 2.13 and Figure 2.14 show the model-predicted survival probability against the fraction actually surviving, plotted for eight different time periods, with each period being the time elapsed between a certain survey wave in the period 1992–2006 and the final wave in 2012. Each dot represents a two-year age bin. We discard age bins with less than 200 observations. The top left graph shows the age bins that were in the panel in the first wave and, as shown by Figure 2.1, there

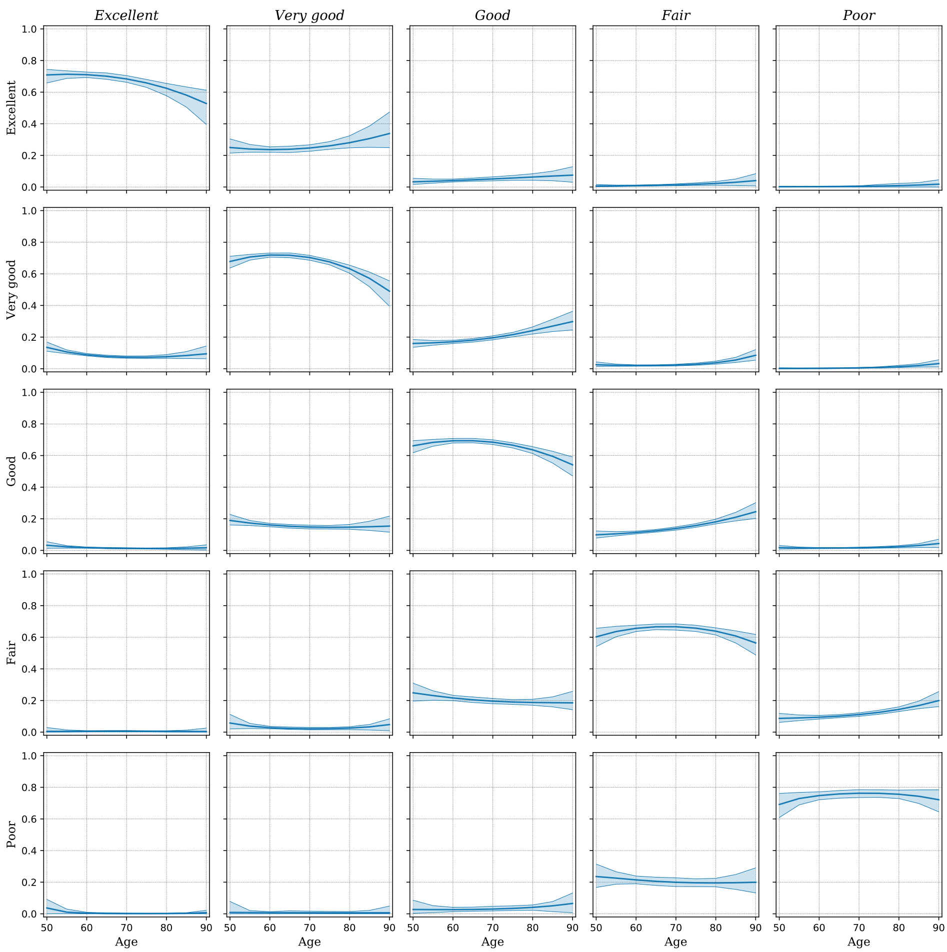


Figure 2.6: One-year health-to-health transition probabilities conditional on survival for nonblack men (model estimates). Initial health states in rows, terminal health states in columns. Shaded areas indicate bootstrapped 95% confidence intervals.

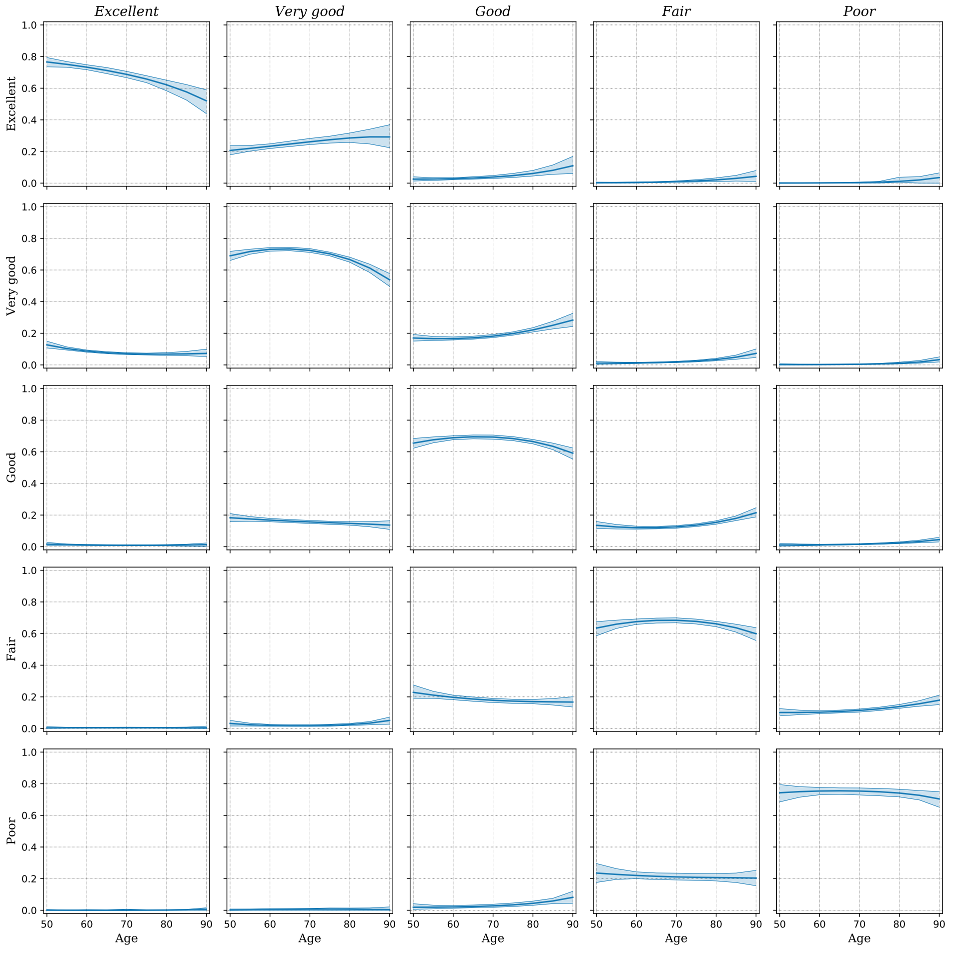
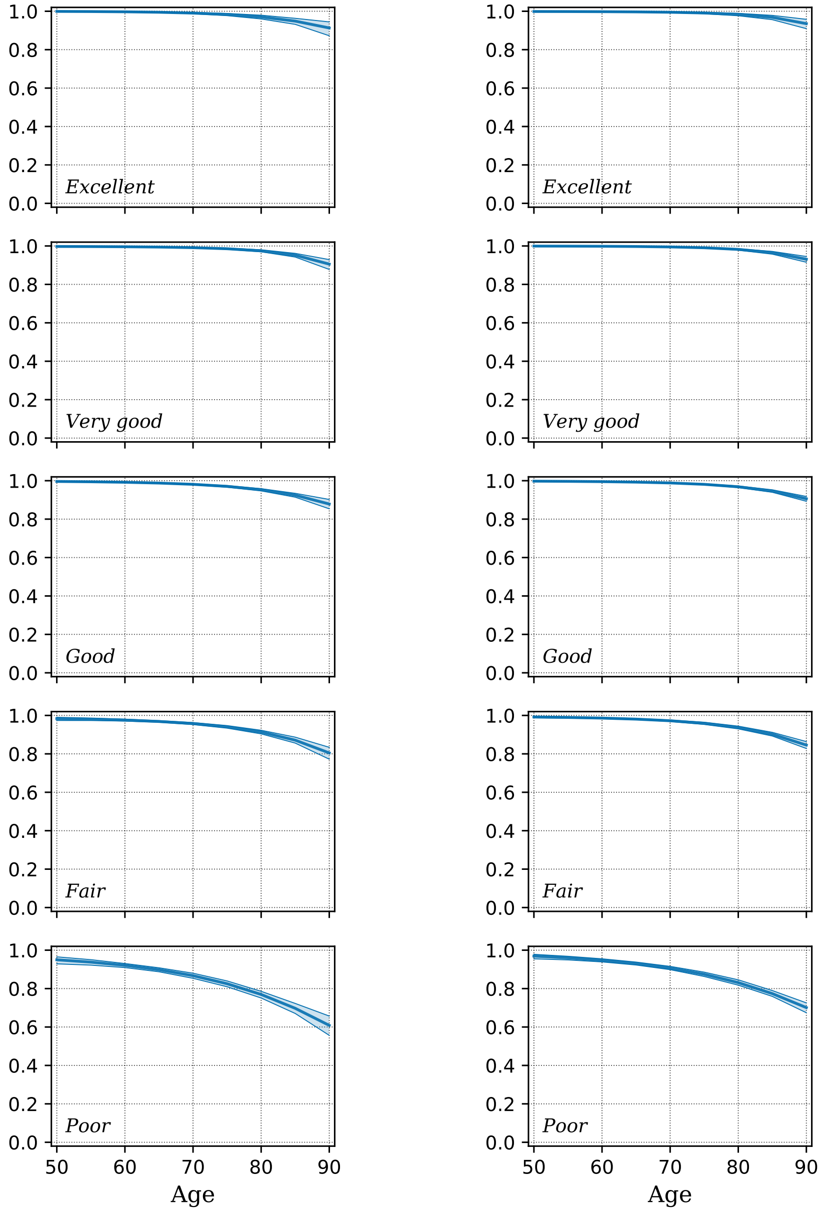


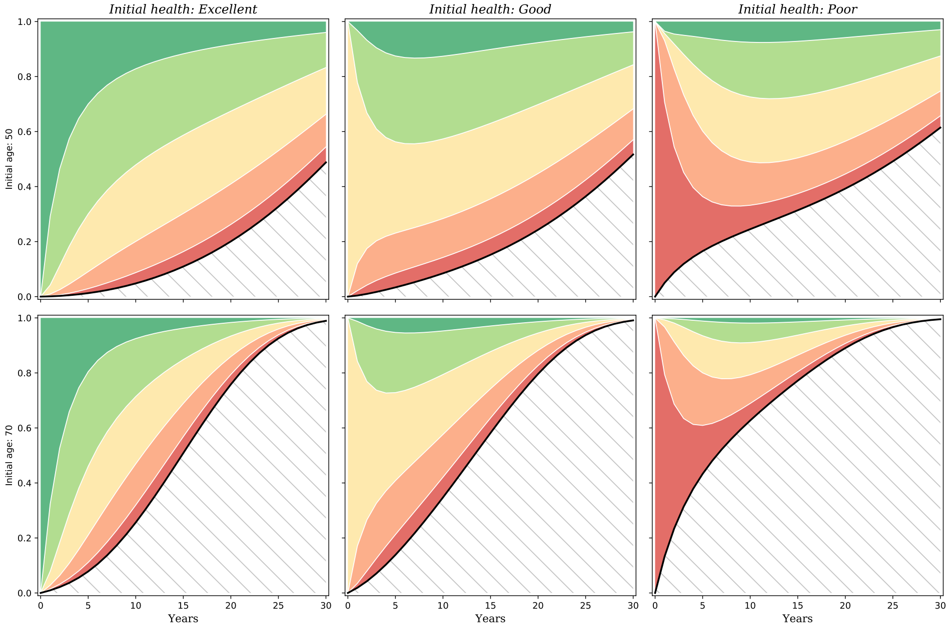
Figure 2.7: One-year health-to-health transition probabilities for nonblack women (model estimates). Initial health states in rows, terminal health states in columns. Shaded areas indicate bootstrapped 95% confidence intervals.



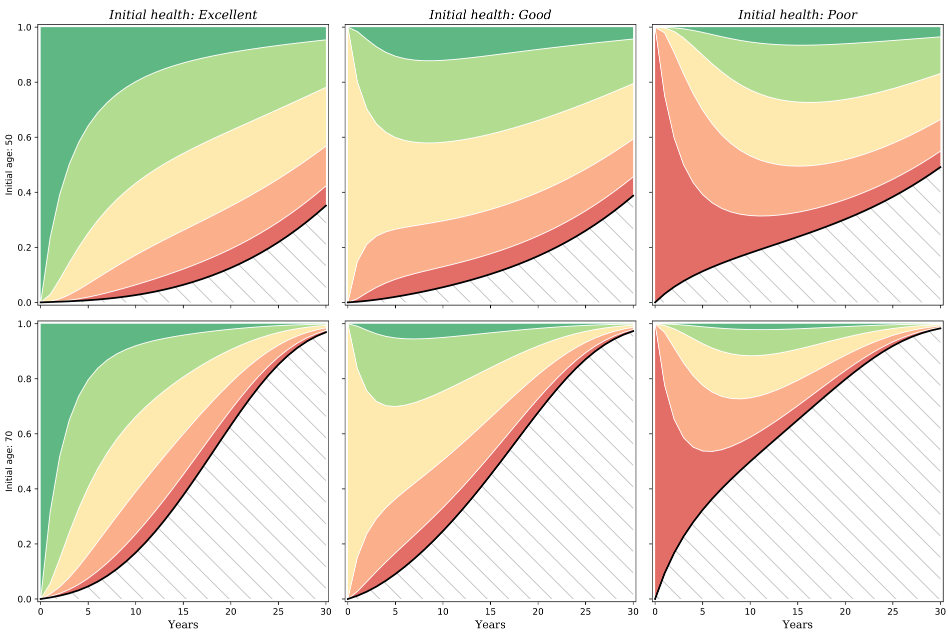
(a) Nonblack men

(b) Nonblack women

Figure 2.8: One-year survival probability (model estimates). Shaded areas indicate bootstrapped 95% confidence intervals.



(a) Nonblack men



(b) Nonblack women

Figure 2.9: Survival probability conditional on initial health state. X-axis indicates years after initial age (upper row 50, lower row 70 years). The colors indicate probability per health state (dark green being the best health state, red the worst). The hatched area represents the probability of being dead.

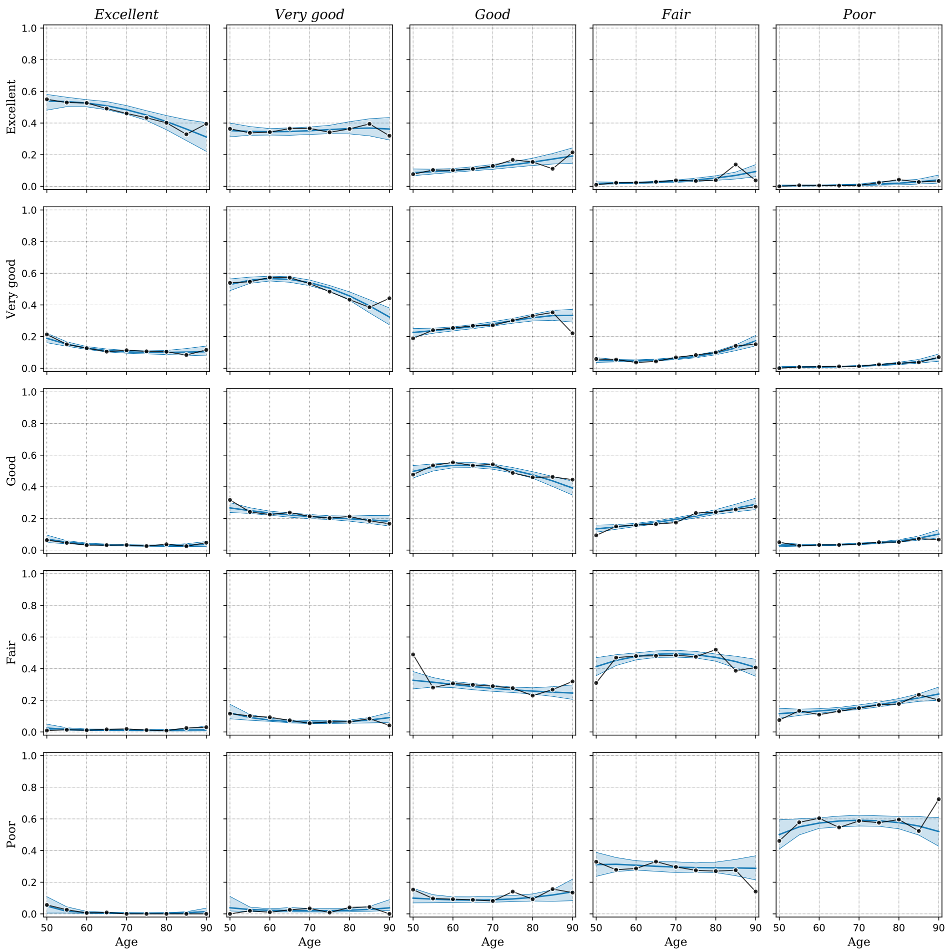


Figure 2.10: Two-year health-to-health transitions conditional on survival for nonblack men: data vs. model. Initial health states in rows, terminal health states in columns. The data includes all two-year observations (i.e., excluding shorter and longer transitions), model predictions are the two-year predictions. Shaded areas indicate bootstrapped 95% confidence intervals.

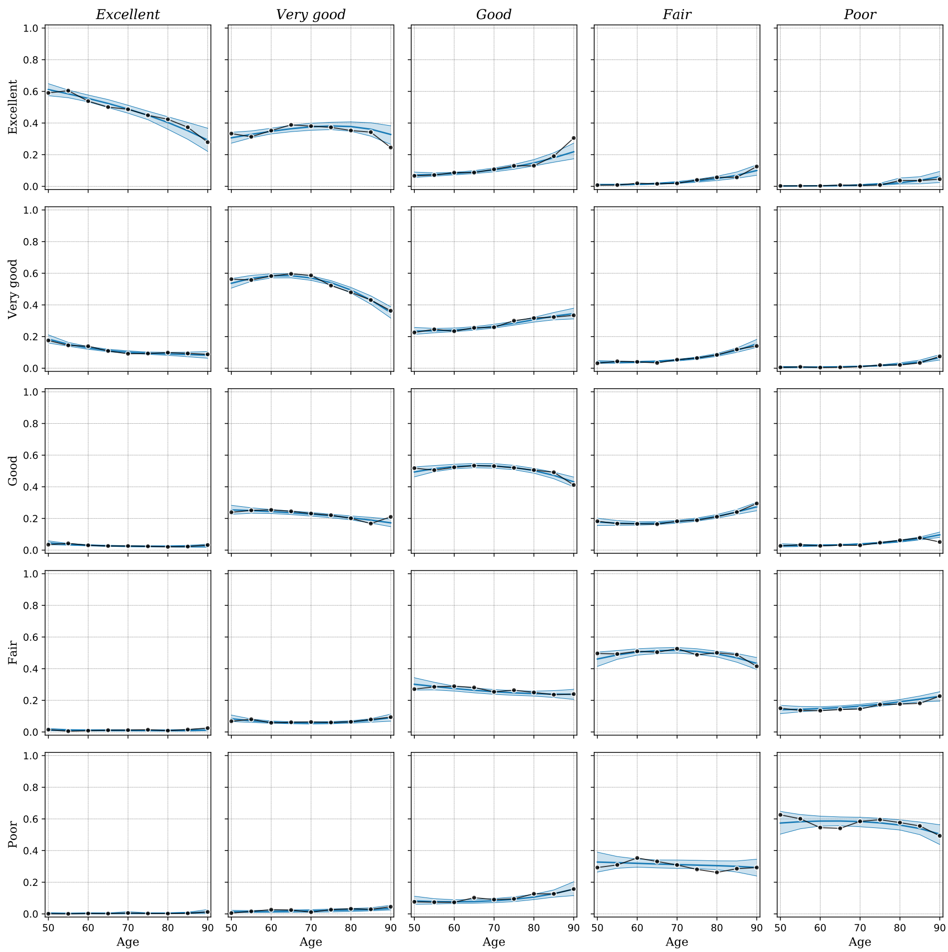
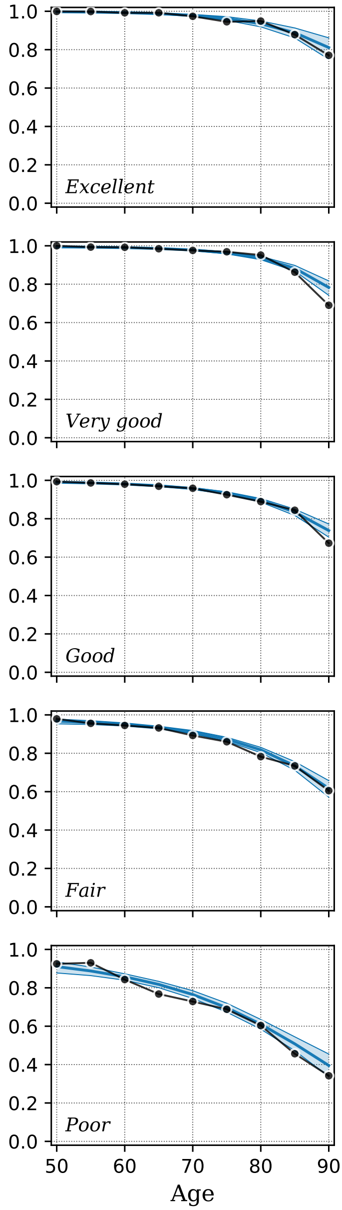
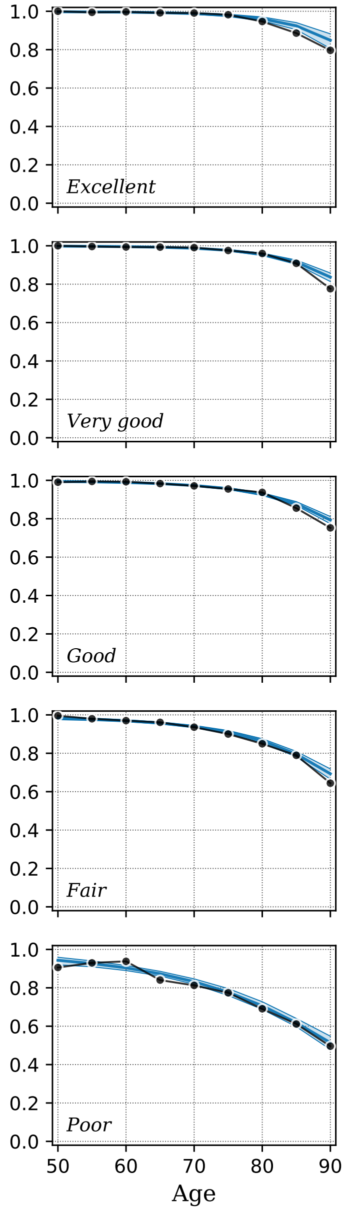


Figure 2.11: Two-year health-to-health transitions conditional on survival for nonblack women: data vs. model. Initial health states in rows, terminal health states in columns. The data includes all two-year observations (i.e., excluding shorter and longer transitions), the model predictions are the two-year predictions. Shaded areas indicate bootstrapped 95% confidence intervals.



(a) Nonblack men



(b) Nonblack women

Figure 2.12: Two-year survival probability: data vs. model. The data includes all deaths within two years after the previous observation, the model prediction is the cumulative probability after two years. Shaded areas indicate bootstrapped 95% confidence intervals.

was only one cohort in the sample at that point. Therefore, there are fewer dots in this graph than in the others. As can be seen, the estimated model captures the survival probability very well.

Life expectancy conditional on health

To calculate the life expectancy conditional on health, we need to take into account all future health-to-health transition probabilities, since the road to death could go via any health transition path. The measure answers the question: what is the life expectancy of a person we observe at age t in health state h? It is computed as follows:

$$e_{h,t} = \left[\sum_{\tau=t}^{T} \sum_{k=1}^{H} \tau \times (1 - p^{s}_{\tau+1|k}) \times \mu_{k,\tau} \right] + \frac{1}{2}$$

where

$$\mu_{j,\tau+1} = \sum_{k=1}^{H} p^{h,j}_{\tau+1|k,s} \times p^{s}_{\tau+1|k} \times \mu_{k,\tau}$$
$$\mu_{j,t} = (0, \ldots, 1, \ldots, 0)$$

where the position of 1 corresponds to the initial health state. The addition of the half year is to correct for the fact that people do not die exactly on their birthday, but deaths are instead approximately uniformly spread out over the year.

Figure 2.15 shows the resulting life expectancies for men and women conditional on the initial health state. As can be seen, the health gradient is substantial: the difference in expected life length between a 50-year-old nonblack man in the best and in the worst health state is 6 years.

Comparing to life tables

The HRS data we use is from the period 1992 to 2012. If we had had many more individuals born each year, we could have computed cohort specific health and survival probabilities by age. However, the sample is not large enough to permit this. Instead, the survival probabilities we calculate should be viewed as *period* life expectancies for the sample period as a whole, and correspond to a weighted average of what is reported in the period life tables by the National Vital Statistics System (NVSS) during those years.[17]

17. There are two types of life tables: period (or current) life tables and cohort (or generation) life tables. The period life table, which is what you find in a regular life table, presents what would happen to a hypothetical cohort if it experienced the mortality conditions of a particular period in time throughout its entire life. The cohort life table, on the other hand, presents the mortality experience of a particular birth cohort from the moment of birth through consecutive ages.

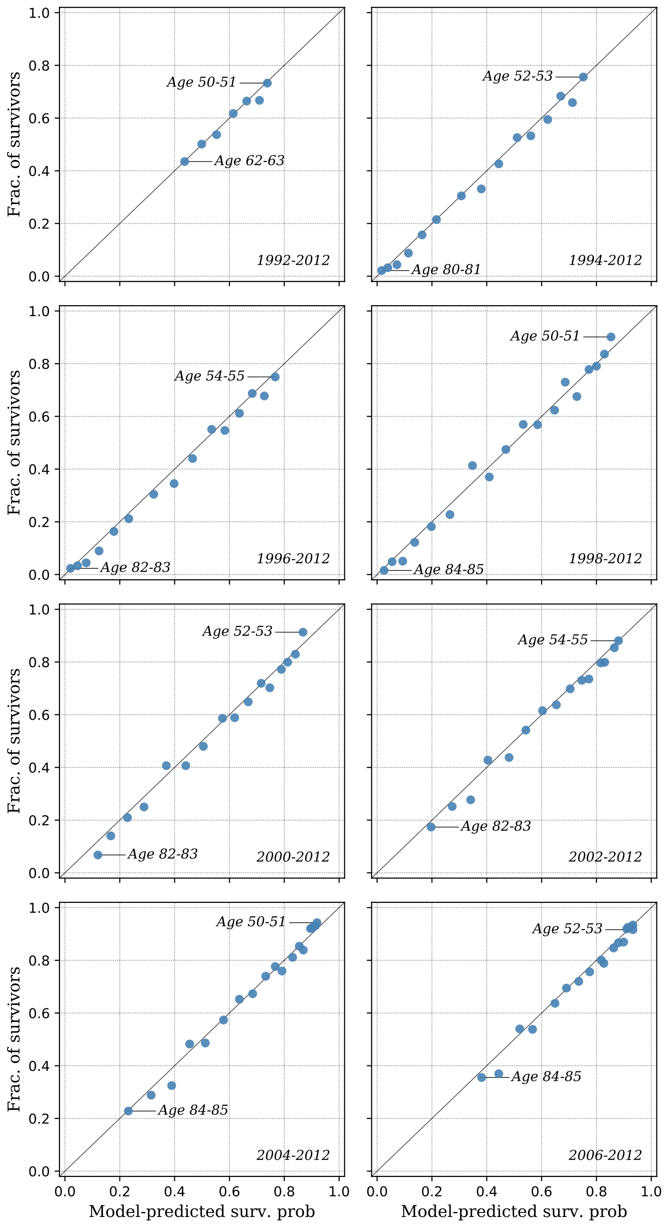


Figure 2.13: Nonblack men: Model-predicted survival probability (on the x-axis) against the fraction actually surviving (on the y-axis), plotted for eight different time periods. The top left graph represents the time period between the first wave (1992) and the last wave (2012). Each dot represents a two-year age bin.

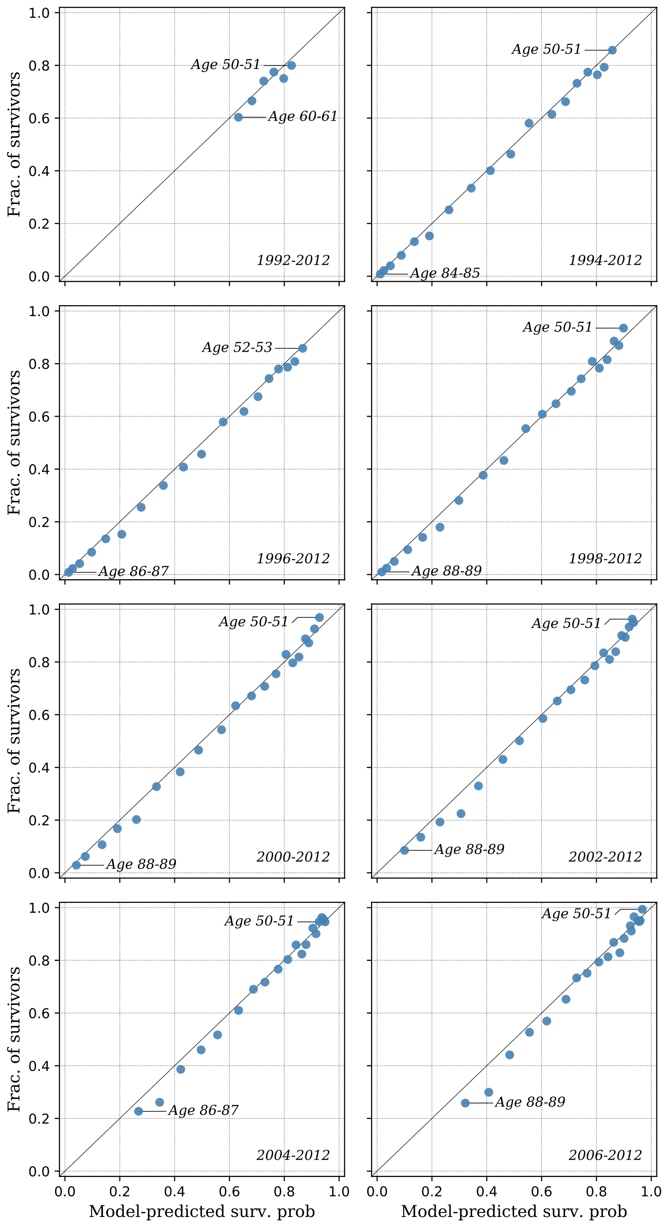
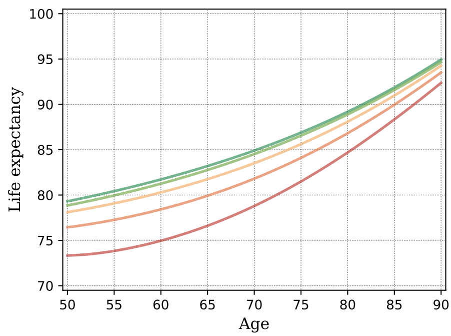
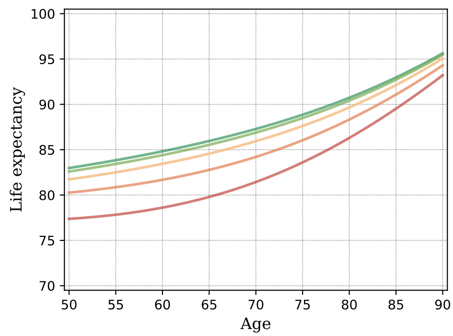


Figure 2.14: Nonblack women: Model-predicted survival probability (on the x-axis) against the fraction actually surviving (on the y-axis), plotted for eight different time periods. The top left graph represents the time period between the first wave (1992) and the last wave (2012). Each dot represents a two-year age bin.



(a) Nonblack men



(b) Nonblack women

Figure 2.15: Life expectancy by age and health state.

We calculate the model-implied life expectancy at the age of 50 given our estimates and given the observed health distribution at this age. Our model gives 78.2 for men and 81.9 for women. This is well in line with what is reported by the NVSS during this period: for white men the NVSS life expectancy at the age of 50 ranges from 77.0 (in 1993) and 79.9 (in 2012) during the sample period. For white women it ranges from 81.7 (in 1993-1996) to 83.4 (in 2012). Hence, the model predictions are within what is reported by NVSS.[18]

The same calculation done for 70-year olds confirms the conclusion. Our model predictions, based on our estimates and the observed health distribution, are 83.3 for nonblack men and 85.7 for nonblack women. The corresponding ages reported by NVSS during the period 1992 to 2012 ranges between 82.3 (in 1993) and 84.4 (in 2012) for men. For women the NVSS reports estimates between 85.3 years (in 1993) and 86.5 years (in 2012). Once more, the model predictions are in line with the national averages reported by the NVSS for the white population.

Duration dependency

It is common in the literature to aggregate the health states into two coarser categories: *good* (covering excellent, very good, and good health) and *bad* (covering fair and poor). However, a benefit of actually using the more fine-grained grid is that it is possible to capture the transition dynamics more precisely.

It has been shown that the probability of transitioning from the coarser *bad* health

18. As pointed out by Pijoan-Mas and Ríos-Rull 2014, life expectancies computed from the HRS should differ slightly from the national average, since the HRS does not include institutionalized individuals.

state into the *good* health state decreases with time: the longer an individual has been unhealthy, the less likely he/she is to become healthy again (De Nardi, Pashchenko, and Porapakkarm 2017). To capture this duration dependency using a health process with only two states, it is necessary to use a higher-order Markov chain. However, using a five-state process, and following the literature by classifying the two worst health states as *bad*, we partly capture this duration dependency even though the process is a first-order only.

As a stylized example, consider the following health-to-health transition matrix:

$$h = \begin{bmatrix} 3/4 & 1/4 & 0 & 0 & 0 \\ 1/4 & 1/2 & 1/4 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/4 \\ 0 & 0 & 0 & 1/4 & 3/4 \end{bmatrix}$$

Assume that all individuals start out at time $t = 0$ in health state 3. We are interested in the individuals in bad health at time $t = 2$, and their probabilities of transitioning back to the good health state, depending on whether they were in bad health for one or two periods.

It turns out that after two periods, 31.25% of individuals are in bad health. 18.75% have been in bad health for two periods, and two thirds of these are in health state 4 while one third are in health state 5. Hence, the probability of transitioning back to good health, conditional on having been in bad health for two periods, is 16.7%.

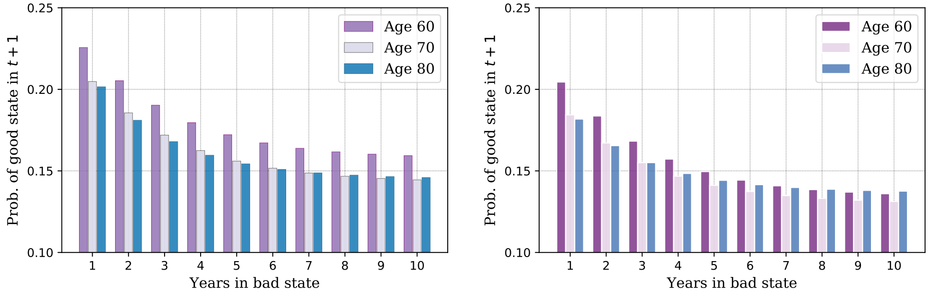
However, the probability of transitioning back to good health for the individuals who have only been in bad health for one period is 25% (this follows immediately since the unhealthy who were in good health in period $t = 1$ can, by construction, only be in health state 4 in this stylized example).

We define a measure of the recovery probability for age t depending on spell length j as:

$$r_t(j) = \Pr\Big(h_{t+1} \in \mathcal{G} \,\Big|\, h_{t-k} \in \mathcal{B} \;\; \forall\, 0 \le k < j,\; h_{t-j} = 3\Big)$$

where $\mathcal{G} = \{1, 2, 3\}$ and $\mathcal{B} = \{4, 5\}$. We choose 3 as the starting health state to simplify the interpretation of the results; the probability of transitioning to a bad health state from health state 1 or 2 is very close to zero, as shown in Figure 2.6 and Figure 2.7.

The results predicted by the full model for $t \in \{60, 70, 80\}$ are shown in Figure 2.16. As can be seen, the probability of recovering from bad health decreases as a function of how many years the individual has spent in bad health. A 60-year old nonblack man who has spent just one year in bad health has a 22.5% probability of recovering, but if he has spent the last five years in bad health the probability is down to 17%. Even though this is not as much duration dependence of bad health as found by



(a) Nonblack men

(b) Nonblack women

Figure 2.16: Probability of recovering from a bad health state, as a function of the number of years spent in bad health.

De Nardi, Pashchenko, and Porapakkarm (2017) in the PSID (they estimate that individuals in the age group 70+ who have spent more than 6 years in bad health have approximately a 12% probability of recovering), it is a substantial improvement of the dynamics compared to a first-order Markov chain with two states, while being more parsimonious than a second-order specification.

2.5 Life expectancy and education level

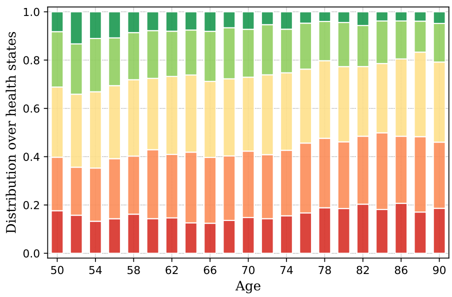
Our model permits the inclusion of time-invariant characteristics to be included in $\boldsymbol{x}_t$, the vector of additional covariates. In this section, we include the level of education in both $g(\bullet)$ and $f(\bullet)$, fully interacted with age and health.[19] We define education in three levels: 1) less than high school, 2) high school, or 3) (some) college or more.[20]

As is well known, more highly educated individuals are healthier. Figure 2.17 shows the distribution of self-reported health for nonblack men for the three education groups. As can be seen, the fraction of individuals in bad health is larger the lower the education.

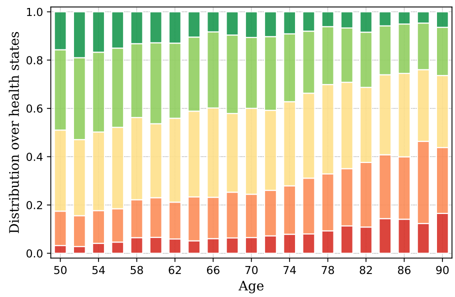
However, even conditional on self-reported health, the health and survival prospects are worse for low-educated individuals. Figure 2.18 shows the evaluation of

19. We also experimented with including a respondent's marital status when first observed, since, e.g., Guner, Kulikova, and Llull 2018 report a health gap between married and unmarried individuals. However, the combination of education and married status makes each group too small to be able to draw any conclusions. Since marital status is not really time-invariant (due to divorce or death of partner) we choose to focus on the education level.

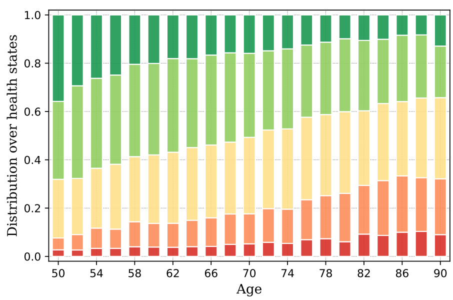
20. In this section, we present the results for nonblack men. Other demographic groups are available upon request.



(a) Less than high school

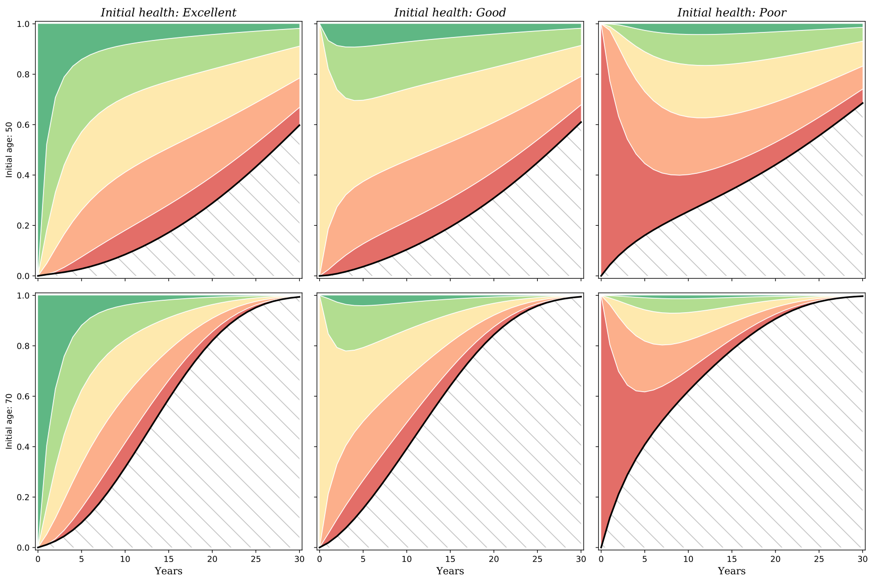


(b) High school

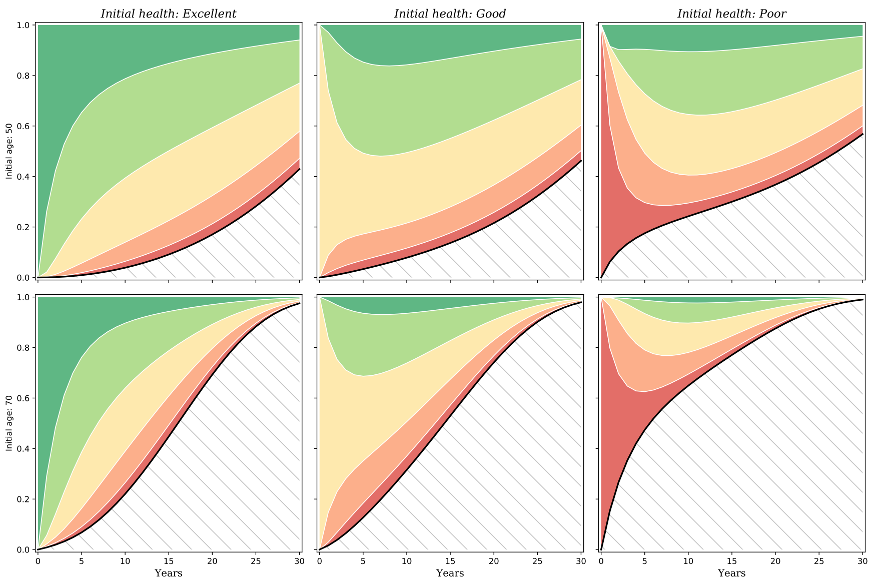


(c) (Some) college

Figure 2.17: Distribution by health state among nonblack men for different ages. Red color indicates worst ("poor") health state while dark green indicates best ("excellent") health state. Observations are grouped into two-year bins.



(a) Less than high school



(b) (Some) college

Figure 2.18: Survival probability conditional on initial health state for nonblack men. The X-axis indicates years after initial age (upper row 50, lower row 70 years). The colors indicate probability per health state (dark green being the best health state, red the worst). The hatched area represents the probability of being dead.

	Using health distr. of		
	No HS	HS	(some) col.
No HS	74.9	75.7	76.0
HS	77.3	78.0	78.3
(some) col.	78.5	79.6	80.1

Table 2.2: Average life expectancy for nonblack men at the age of 50, by education level and initial health distribution. See the main text for further description.

probabilities for nonblack men for each health state and for being dead, given an initial health state and a starting age, contrasting the lowest and the highest education groups.[21]

As shown in the figure, the less educated group is both more likely to transition to bad health, and bad health is more persistent for the less educated group (visibly shown by the fact that the red areas are larger in the graphs for the lower education level). For instance, a 50-year old man with less than high school education, starting out in poor health, has a probability of approximately 30% to survive an additional 30 years. However, for a 50-year old man with (some) college education, but in the same poor initial health, the probability is more than 40%. As another example, a 70-year old in good initial health condition and with less than high school education has a probability of approximately 60% to survive an additional 10 years. For a man of the same age and with the same initial health condition, but with (some) college education, the probability is 70%.

Hence, the life expectancy differs by education level. Figure 2.19 shows the life expectancy by education level, age and health state and, as can be seen, the life expectancy, conditional on health state, is higher for more highly educated people.

The *average* life expectancy for 50-year old nonblack men with less than high school education is 74.9 years, while the average for 50-year olds with (some) college education is 80.1 years. This difference is a result of two factors. First, the group with lower education has worse average self-reported health at the initial age of 50. Second, conditional on health, their health dynamics and survival probabilities are worse from this age and onwards.

To disentangle these two effects we make the following experiment: we take the health and survival process of the lowest education group but use the initial health distribution of the highest education group. The average life expectancy for this hypothetical group of 50-year olds would be 76.0 years. Hence, the difference in life expectancy for high vs low educated 50-year old men is approximately one fifth due

21. Results for the second group, with high school education, are shown in the appendix, as well as the full health transition matrices and the survival probabilities by education level.

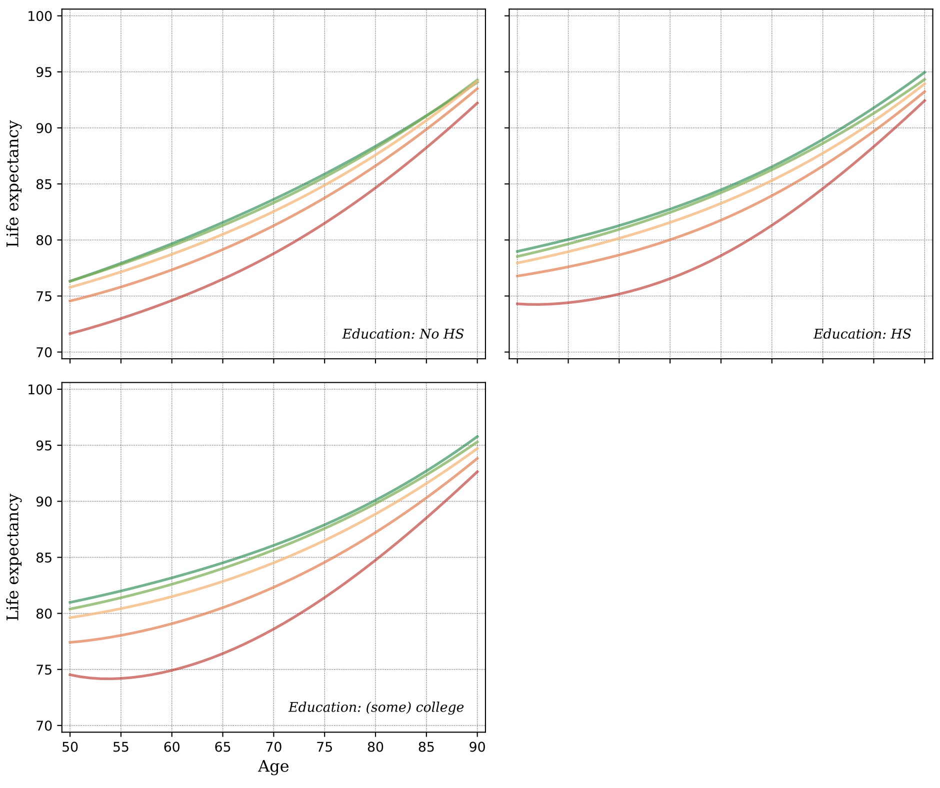


Figure 2.19: Life expectancy by education level, age, and health state for nonblack men.

to worse overall initial health, and four fifths due to worse health dynamics after that age. Table 2.2 shows the full set of combinations of health and survival processes and initial health distributions. As can be seen, it is generally true that the health and survival process after the age of 50 has a larger effect on life expectancy than the health distribution at the age of 50.

2.6 Conclusion

Many studies have identified health dynamics and health shocks as a major source of risk faced by individuals. To incorporate this risk in life-cycle models, we need a health and survival process that captures the main features of the data, while still keeping the formulation parsimonious. In this paper, we provide improved estimates for annual age-dependent health transitions and survival probabilities for different subsamples of the U.S. population.

The resulting Markov process can be used in any life-cycle model where the risks associated with health and survival are of interest. The estimated health process shows high persistence and the survival probability differs substantially depending on health. The estimated difference in expected longevity for a 50-year-old white man in excellent vs. poor health is 6 years.

The estimation method we propose is specifically designed to handle the peculiarities of the health and death data in the HRS. However, besides the specificities of the absorbing state of death, the method for estimating age-dependent one-year transition probabilities from data that is observed at irregular and overlapping intervals could be used outside the health and survival realm. It could be applied to any process where the dependent variable is ordinal (or limited in other ways).

Bibliography

Amengual, Dante, Jesús Bueren, and J Crego. 2017. *Endogenous health groups and heterogeneous dynamics of the elderly.* Technical report. HEDG, c/o Department of Economics, University of York.

Arias, Elizabeth. 2014. United States life tables, 2009. *National vital statistics reports,* Working Paper Series, 62 (7).

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. 2016. The association between income and life expectancy in the united states, 2001-2014. *Jama* 315 (16): 1750–1766.

Coile, Courtney, Kevin S. Milligan, and David A. Wise. 2016. *Health capacity to work at older ages: evidence from the U.S.* Working Paper, Working Paper Series 21940. National Bureau of Economic Research.

Contoyannis, Paul, Andrew M. Jones, and Nigel Rice. 2004. The dynamics of health in the british household panel survey. *Journal of Applied Econometrics* 19 (4): 473–503.

De Nardi, Mariacristina, Eric French, and John B. Jones. 2010. Why do the elderly save? the role of medical expenses. *Journal of Political Economy* 118 (1): 39–75.

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. *The lifetime costs of bad health.* Technical report. National Bureau of Economic Research.

Finkelstein, Amy, Erzo F. P. Luttmer, and Matthew J. Notowidigdo. 2013. What good is wealth without health? The effect of health on the marginal utility of consumption. *Journal of the European Economic Association* 11:221–258.

French, Eric. 2005. The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies* 72 (2): 395–427.

French, Eric, and John Bailey Jones. 2011. The effects of health insurance and self-insurance on retirement behavior. *Econometrica* 79 (3): 693–732.

Guner, Nezih, Yuliya Kulikova, and Joan Llull. 2018. Reprint of: marriage and health: selection, protection, and assortative mating. Gender Differences in the Labor Market, *European Economic Review* 109:162–190.

Hosseini, Roozbeh, Kai Zhao, Karen Kopecky, et al. 2018. The evolution of health over the life cycle. In *2018 meeting papers.* 977. Society for Economic Dynamics.

Idler, Ellen L., and Yael Benyamini. 1997. Self-rated health and mortality: a review of twenty-seven community studies. *Journal of Health and Social Behavior* 38 (1): 21–37.

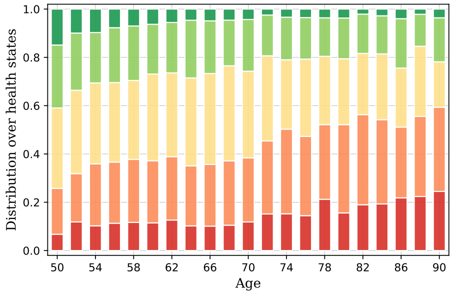
Kopecky, Karen A., and Tatyana Koreshkova. 2014. The impact of medical and nursing home expenses on savings. *American Economic Journal: Macroeconomics* 6 (3): 29–72.

Pijoan-Mas, Josep, and José-Víctor Ríos-Rull. 2014. Heterogeneity in expected longevities. *Demography* 51 (6): 2075–2102.

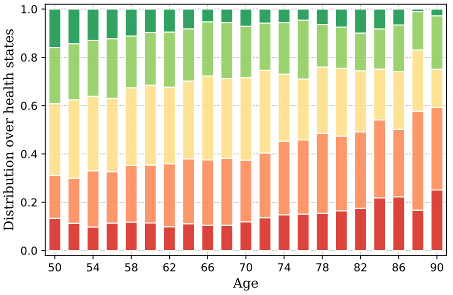
Appendix

2.A Health distribution

The health distribution by age for black women and black men is shown in Figure 2.20.



(a) Black women

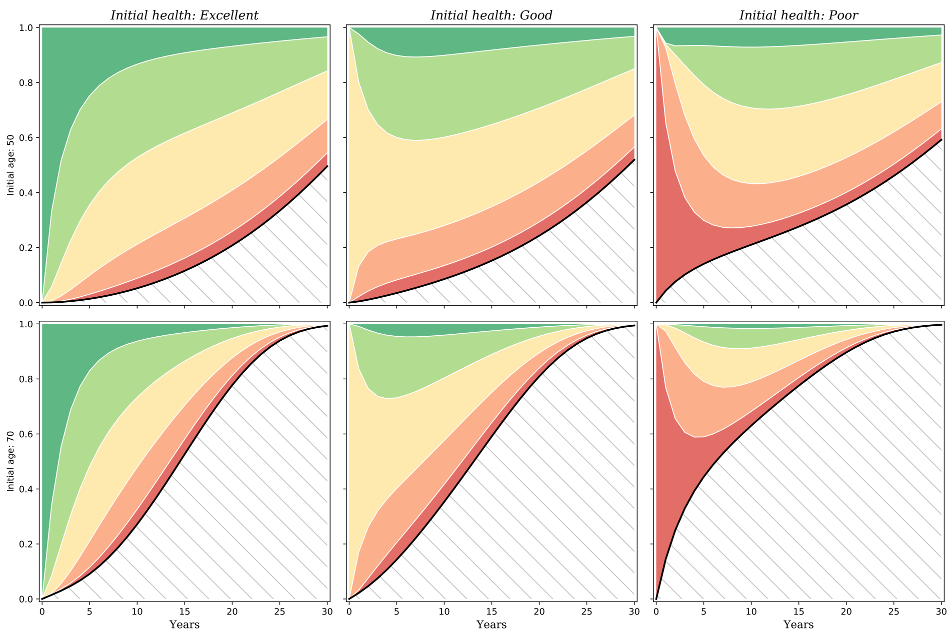


(b) Black men

Figure 2.20: Distribution by health state for different ages. Red color indicates worst ("poor") health state while dark green indicates best ("excellent") health state. Observations are grouped into two-year bins.

2.B Evaluation of probabilities for health state: education group 2

Evaluation of probabilities for each health state and for being dead, given an initial health state and a starting age, for nonblack men with high school education.



(a) High school

Figure 2.21: Survival probability conditional on initial health state for nonblack men. The x-axis indicates years after initial age (upper row 50, lower row 70 years). The colors indicate probability per health state (dark green being the best health state, red the worst). The hatched area represents the probability of being dead.

2.C Health transition matrices and survival probabilities by education level

Figure 2.22, Figure 2.23, and Figure 2.24 show the health transition matrices for the different education levels. Figure 2.25 shows the resulting survival probabilities, conditional on health state, age, and education level.

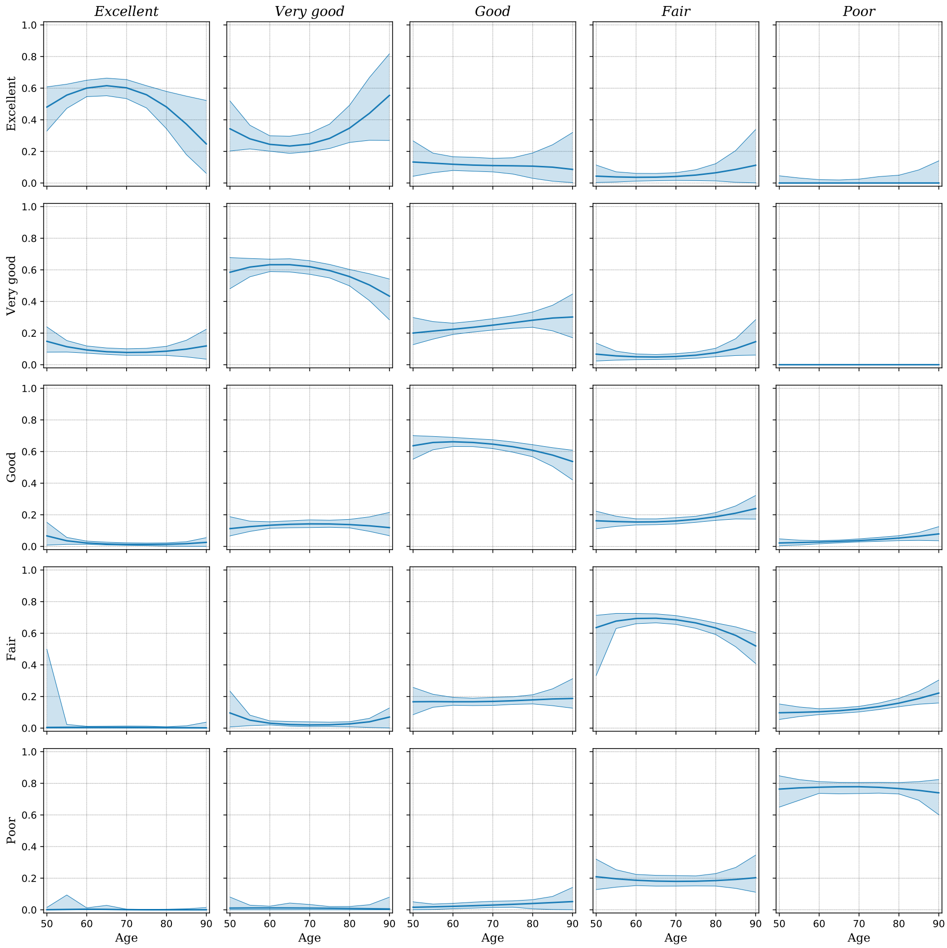


Figure 2.22: One-year health-to-health transition probabilities for nonblack men with *less than high school* (model estimates). Initial health states in rows, terminal health states in columns. Shaded areas indicate bootstrapped 95% confidence intervals.

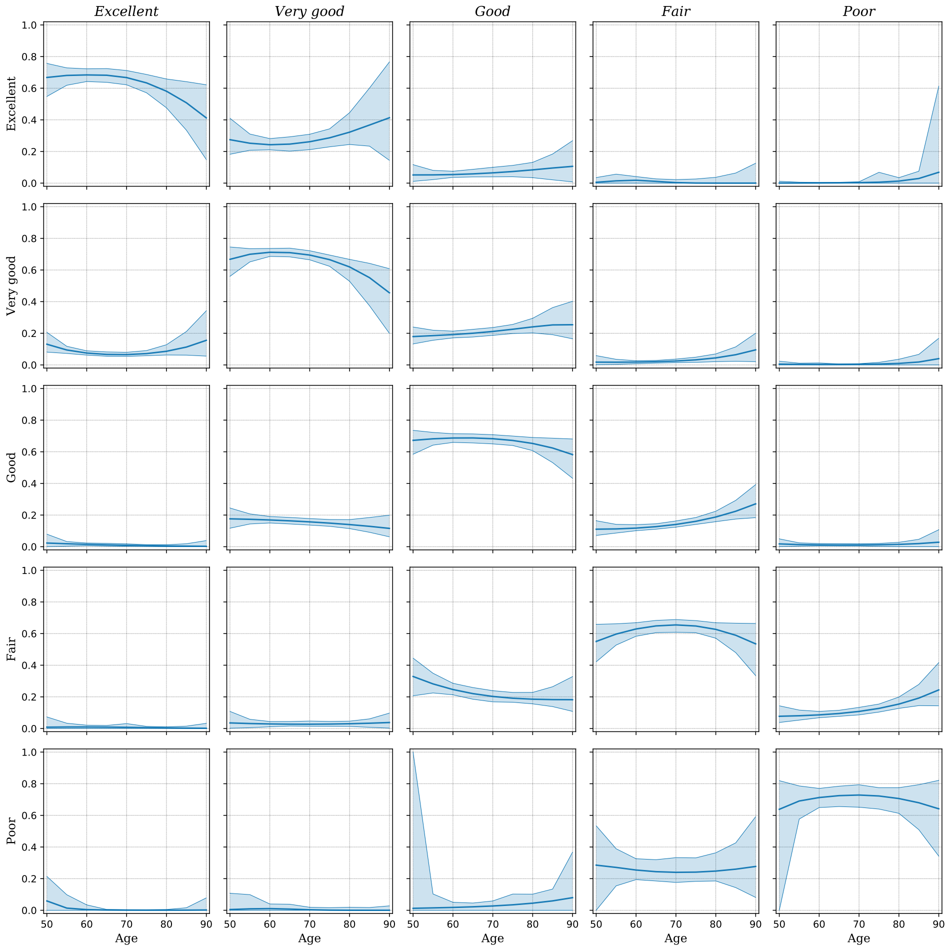


Figure 2.23: One-year health-to-health transition probabilities for nonblack men with *high school* (model estimates). Initial health states in rows, terminal health states in columns. Shaded areas indicate bootstrapped 95% confidence intervals.

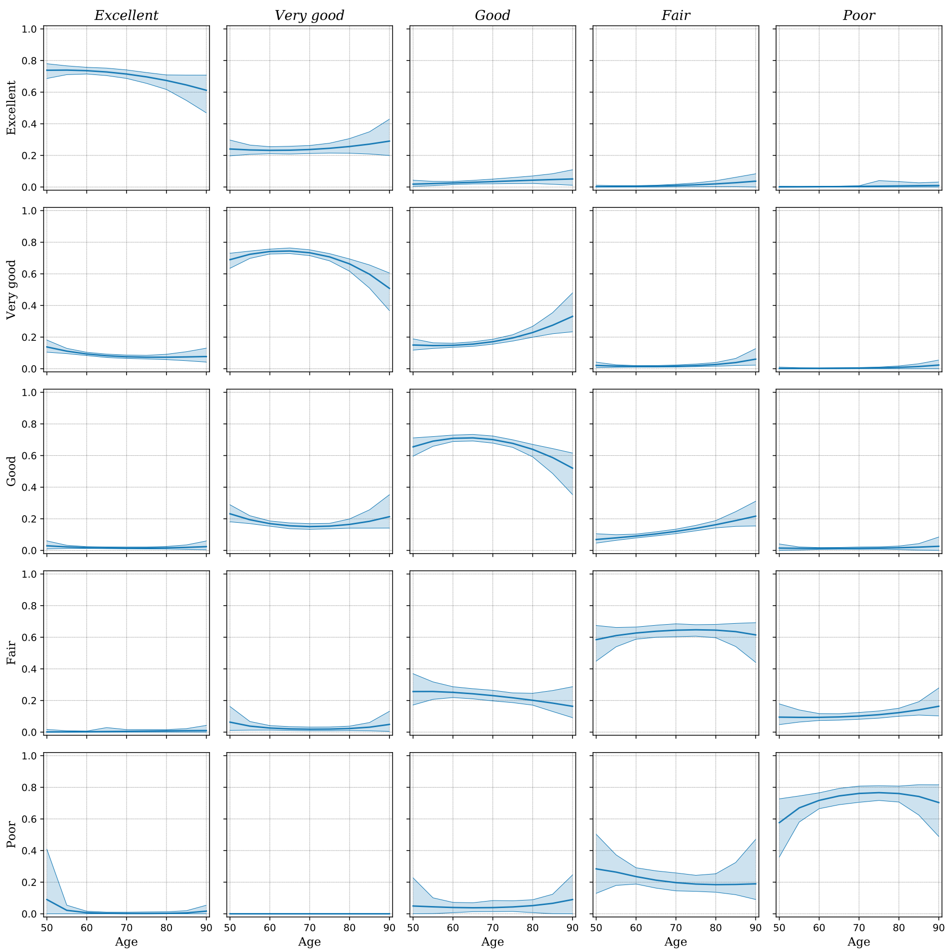
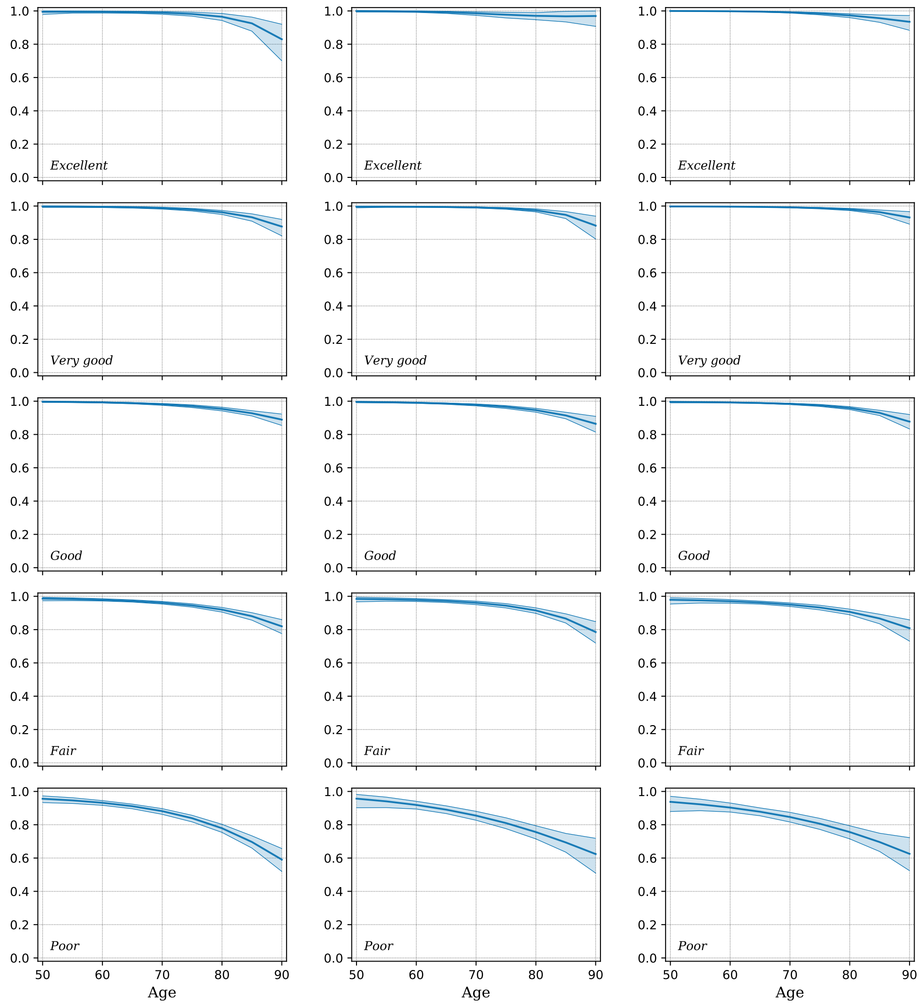


Figure 2.24: One-year health-to-health transition probabilities for nonblack men with *(some) college* (model estimates). Initial health states in rows, terminal health states in columns. Shaded areas indicate bootstrapped 95% confidence intervals.



(a) Less than high school **(b)** High school **(c)** (Some) college or more

Figure 2.25: One-year survival probability (model estimates) for nonblack men, by education level. Shaded areas indicate bootstrapped 95% confidence intervals.

Chapter 3

Subjective Life Expectancies, Time Preference Heterogeneity and Wealth Inequality

(joint with Jonna Olsson)

3.1 Introduction

There is a substantial heterogeneity in life expectancy in the population: an average 70-year-old man in excellent health has a 75% chance of reaching his 80th birthday, while the corresponding probability for a man in poor health is just below 40%. According to standard economic theory, the healthy man should save more for the future, given the higher probability of living a longer life. The question we ask in this paper is how heterogeneity in life expectancy affects savings rates and ultimately wealth inequality – within a cohort and in the aggregate economy.

Preference heterogeneity, and especially heterogeneity in time preferences, is one of the potential sources of wealth inequality that have been used in previous literature, but preferences are difficult to measure and quantify.[1] In this paper, we investigate one type of time preference heterogeneity, namely heterogeneity in life expectancy, which we document using micro data.

An individual's consumption/savings decision is not necessarily guided by the objective statistical life expectancy, but rather by the individual's beliefs about survival. We document new facts about a systematic *bias* in these beliefs: individuals with a low survival probability relative to their peers underestimate, while individuals with a high survival probability overestimate, their life expectancies. This systematic bias exacerbates the survival expectancy heterogeneity in the population.

To gauge the effect of survival heterogeneity on inequality, we use an overlapping-generations general-equilibrium model with uninsurable idiosyncratic shocks. Agents face heterogeneous survival risk that depends on their current health state, and are

1. See De Nardi and Fella 2017 for further references.

subject to health shocks that follow a process estimated from data. Besides this uncertainty, we also include standard persistent and transitory shocks to labor productivity. Since we are interested in savings behavior in late life, it is important to capture other incomes during this period. Therefore we carefully model retirement benefits, closely mimicking the U.S. social security system.

We compare several scenarios along two main dimensions: first, we vary the model environment according to how agents form expectations about survival, which can be uniform survival beliefs (without health heterogeneity), objective survival beliefs, or subjective survival beliefs. In the latter two, survival beliefs are allowed to depend on an individual's health. As for the second dimension, we examine economies in which agents either do or do not have a warm-glow bequest motive, since this turns out to be central to household behavior. Together, these dimensions give rise to six different model environments, where the model without survival belief heterogeneity or a bequest motive serves as the benchmark case.

We show that the standard life-cycle model gives rise to counter-factual implications when introducing survival heterogeneity. In an environment without bequests, agents with a longer life expectancy save more, as expected. This is in line with the data, where individuals in better health have larger asset holdings. The effect on within-cohort inequality at higher ages is substantial, but the overall effect on wealth inequality in the economy is small. The reason is that the richest individuals are those of ages 60–64. At those relatively younger ages, the rich, who are also the healthiest, have a life expectancy which is not too far from the average and therefore their savings rate is only marginally affected. Moreover, their subjective beliefs are close to the objective survival probabilities, and therefore the additional effect of subjective beliefs is small.

However, as is well known, this standard model without a bequest motive gives rise to counterfactually low savings among the elderly since agents draw down their assets to virtually zero late in life.[2] In the data, on the other hand, individuals do, on average, have substantial asset holdings even beyond the age of 80. Therefore, we add a warm-glow bequest motive as in De Nardi 2004, which is the most commonly used formulation in the macroeconomic literature.

We calibrate the bequest parameters so that the benchmark scenario without survival heterogeneity approximately matches the median asset holdings at older ages observed in the data. The effect of introducing survival heterogeneity into this environment is counter-intuitive and perhaps also unexpected: agents in poor health now save *more* than their healthy counterparts. The reason is as follows: since agents in poor health are more likely to die soon, they put an increased weight on bequest

2. In a model without retirement benefits, agents keep slightly more assets to be able to sustain consumption in the case of an unexpectedly long life. However, even then, asset holdings are far lower than what we observe in the data.

utility, thus raising their incentives to save. Hence, there are two effects from lower life expectancy that work in opposite directions: a shorter expected life span makes the agent save less for own consumption, but a stronger bequest motive makes the agent save more. The net effect varies depending on the calibration of bequest parameters, but the second mechanism is always present with a bequest formulation of this type: a shorter life span makes agents want to save more to leave bequests. This creates a health-wealth gradient that is counter-factual and we argue that this mechanism is implausible.

We conclude that none of the standard models are adequate for investigating the effect of survival heterogeneity on savings rates and wealth inequality. We discuss possible extensions and reformulations and point out directions for further research.

This paper speaks to three broad strands of literature. The first is macroeconomic studies pointing out the importance of heterogeneity in time preferences to explain wealth inequality (Krusell and Smith 1998; Hendricks 2007). Compared to these studies, we analyze the importance of a preference heterogeneity that is micro-founded by health heterogeneity.

The second is the literature about the general impact of health on wealth (De Nardi, Pashchenko, and Porapakkarm 2017; Poterba, Venti, and Wise 2017; Lee and Kim 2008; Coile and Milligan 2009; Smith 1999). These studies incorporate multiple links between health and economic outcomes, while we restrict ourselves to a very specific channel, namely survival heterogeneity. We argue that this channel is interesting in itself, and that it is necessary to understand it correctly in order to include it in the broader assessment of the health-wealth gradient.

The third is the literature concerned with subjective survival expectations (Ludwig and Zimper 2013; Groneck, Ludwig, and Zimper 2016, 2017; Heimer, Myrseth, and Schoenle 2019; Hurd and McGarry 2002; Hamermesh 1985; Smith, Taylor, and Sloan 2001). Many studies have documented the existence of an age bias in subjective life expectancies, and a few of the papers within this group are concerned with the implications for the consumption/savings behavior. However, none have documented the within-cohort bias, or analyzed the implications for wealth inequality.

In the next section, we briefly describe how we estimate the health and death process and give details about the systematic bias in survival expectations in our sample. Section three describes the model we use to quantify the importance of the heterogeneity in survival expectations. After that, we discuss the parametrization and present our results. The last section concludes the paper and comments on the implications of these findings, and points out directions for further research.

3.2 Empirical evidence

Data

We use the Health and Retirement Study (HRS), a representative panel of elderly U.S. households, to investigate the evolution of health and longevity in the later stages of life. The survey includes, among other things, questions about self-reported health, expectations about survival, and date of death, if applicable.

The survey started in 1992 and in this paper, we use HRS data up to and including the eleventh wave in the year 2012.[3] The first cohort included in the survey was between 51 and 61 years old in 1992, and thereafter new cohorts have been added. Many of the respondents have died over the sample period, making it an appropriate data set for studying survival.

The health-wealth gradient

Figure 3.1 shows net total wealth over the life cycle, by self-reported health status. The health–wealth gradient is well documented, but the underlying causal relationship is debated. One line of argument, especially common in the medical discipline, is that low economic status leads to poor health. The underlying reasons could be many: poor people have access to less or lower quality medical care, do not invest enough in preventive health measures, and/or have more health-deteriorating habits.

However, there are also many arguments for the reversed causality: poor health has economic consequences in itself. First, poor health may restrict the individual's earnings potential by making it more costly to work and/or lowering the wage. Second, poor health may lead to large medical expenditures. Third, poor health may lower the savings incentives due to a lower survival expectancy.

In this paper, we focus on the savings incentives channel. Our aim is to answer the question: how much of the difference in wealth accumulation between individuals in good health and bad health is due to their (perceived) difference in longevity? We go beyond the objective survival expectations and also look at how the consumption/savings decision is affected by an individual's *subjective* survival expectations.

There are a few empirical studies that corroborate the existence of this channel and suggest a causal link. For instance, Heimer, Myrseth, and Schoenle 2019 show that greater survival optimism correlates with higher savings rates, not only controlling for standard demographic characteristics such education and marital status, as well as income, but also characteristics such as financial literacy and risk tolerance.

3. The RAND version O, covering waves up until 2012, is the most recent RAND release that includes data from the National Death Index (NDI). Since correct death dates are crucial for our analysis, this is our preferred data set. There is one later RAND release, covering the 2014 wave as well, but there NDI data is lacking. An analysis shows that there are discrepancies in death dates between the exit interview information and the NDI date of death.

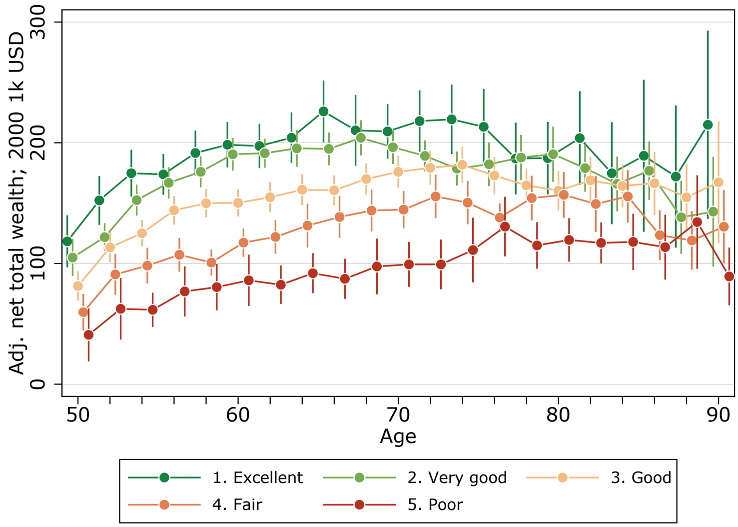


Figure 3.1: Average net total household wealth by self-reported health status (men). Pooled sample from HRS 1992-2012. Assets adjusted for inflation, outliers, business cycles and growth and cohort effects.

Hurd, Smith, and Zissimopoulos 2004 show that individuals with very low subjective survival probabilities retire and claim social security benefits earlier.

In this paper, we examine the isolated effect of heterogeneity in survival expectancy on savings behavior and its implications for wealth inequality. Therefore, we need to formulate heterogeneity in survival expectations, both objective and subjective, in a way that can be used in a structural model. This is what we turn to in the next section.

Objective health and survival probabilities

The HRS contains self-reported health states and a respondent's death date, if applicable. For our model, we need a yearly Markov process for health transitions and survival as a function of the model's state variables age and health. We estimate this Markov process as described in Foltyn and Olsson (2019). Conceptually, the method is a straightforward maximum likelihood estimator, where the probability of observing the transition paths in the data is maximized.

To put structure on the Markov process, we follow Pijoan-Mas and Ríos-Rull 2014 and use a nested logit model, where survival and health transitions conditional on survival are modeled as functions of the current health state and age. The probabil-

ity of survival follows the usual binary-outcome logit model while, conditional on survival, health transitions are modeled using multinomial logit. Thus, it is assumed that the one-period-ahead probability of survival is given by

$$p^s_{t+1} = \frac{1}{1 + e^{-g(h_t, \boldsymbol{x}_t|\gamma)}} \tag{3.1}$$

where h_t is the current health state and the vector $\boldsymbol{x}_t$ contains any other variable of interest, in particular age.

Survival probabilities are governed by the parameter vector $\boldsymbol{\gamma}$ to be estimated (together with $\boldsymbol{\beta}$, a vector governing the health transition probabilities). The result of the estimation of the objective health and survival probabilities is a first-order Markov process for health and the absorbing state of death. This process can be used to calculate objective life expectancies, conditional on age, health, race, and gender. Similarly, we estimate a new set of "subjective parameters" $\widetilde{\boldsymbol{\gamma}}$ that govern subjective beliefs about survival, which we discuss in detail in the next section.

Figure 3.2 illustrates the dynamics of the estimated health and survival process. The figure shows the evaluation of probabilities of each health state and of being dead along a 30-year forecast horizon for a given initial health and age. As can be seen, the survival probability differs substantially depending on the initial health state: for a nonblack man aged 70 in excellent health, the predicted probability of surviving an additional 10 years is more than 80%, while the probability is just around 40% if instead starting out in poor health.

Average expectation errors in survival probabilities

In the expectations survey module of the HRS, respondents are asked about the probability they assign to certain events. One of these questions is about the probability of surviving to a certain age, for example: "Using a number from 0 to 100, what do you think are the chances that you will live to be at least 100 years?"[4]

The target age the respondent is asked about depends on his/her age. For instance, in 1995, respondents below the age of 70 were asked about the probability of living until the age of 80, while respondents above the age of 85 were asked about the target age of 100. In later surveys and for some age intervals, the respondent is asked about several target ages.

Using these elicited beliefs, we compare the average probability that individuals of a certain age assign to survival until a given target age to the probability according to official lifetables. The results are shown in Figure 3.3. As can be seen, there is a systematic error along the age gradient: younger individuals on average tend to

4. Before the respondent answers the questions about expectations, the interviewer discusses probabilities and verifies that the respondent understands the concept.

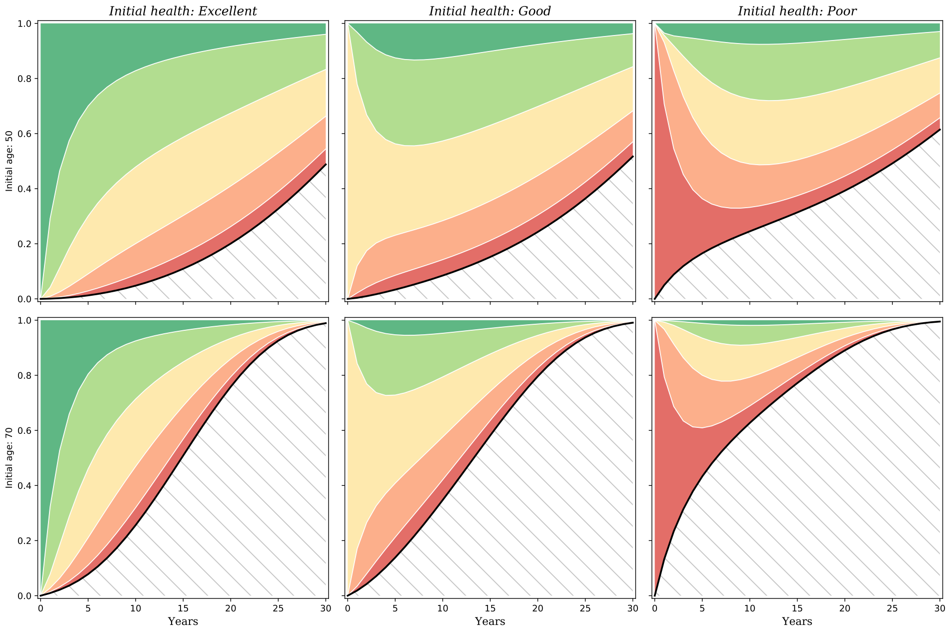


Figure 3.2: Survival probability conditional on the initial health state. The X-axis indicates years after initial age (upper row 50, lower row 70 years). The colors indicate probability per health state (dark green being the best health state, red the worst). The hatched area represents the probability of being dead.

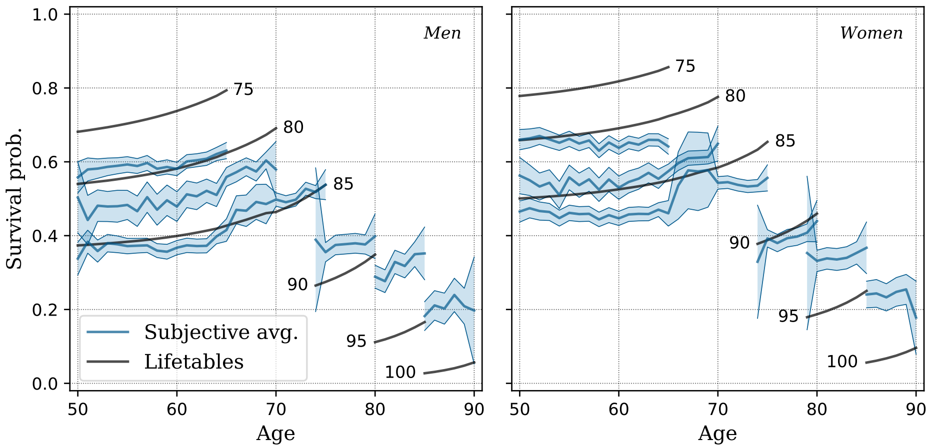


Figure 3.3: Objective vs. subjective survival probabilities as a function of age. The number next to the black line indicates target age. The blue line shows average expectation among the population. Shaded areas indicate bootstrapped 95% confidence intervals.

underestimate, while older individuals tend to overestimate, their survival probability as compared to objective lifetable estimates.

This age-dependent error is a stylized fact in the literature about survival expectations (see, e.g., Ludwig and Zimper (2013), Groneck, Ludwig, and Zimper (2016), and Heimer, Myrseth, and Schoenle (2019) and the references therein). The pattern has been used, e.g., to improve the fit of the asset profile of the canonical life-cycle model with the data: due to underestimation early in life, young agents do not accumulate as much assets, while overestimation in later years dampens the rate at which assets are decumulated.

In the previous section, we did not only estimate the survival probability as a function of age and gender, but also as a function of health and race, which allows us to disaggregate the expectation errors further. Figure 3.4 shows expectation errors by gender, race, age, and health among respondents who have answered the question about their perceived probability of survival until the age of 75. The first observation is the striking positive correlation between subjective self-reported survival probability and the predicted (objective) survival probability, which means that subjective beliefs are informative and not just random noise.[5] The second observation is the systematic

5. This confirms the findings by Smith, Taylor, and Sloan 2001 who compare subjective expectations to actual longevity on an individual basis (with a much smaller sample) and find that expectations and actual survival are consistent, and moreover that expectations are updated in the event of bad health shocks.

bias in beliefs.

As was shown in Figure 3.3, on average individuals underestimate their probability of survival until the age of 75. From looking at Figure 3.4 and focusing on nonblack males, it becomes clear that it is the individuals in bad health that are underestimating their survival probability, while the individuals in excellent health are on average reasonably close to their objective survival probability.

Figure 3.5 and Figure 3.6 show the same information for target ages 85 and 95, respectively. Again, we see that on average, individuals in bad health are more pessimistic than those in good health. For example, as we saw in Figure 3.3, on average individuals overestimate their survival probability when they are asked about a target age of 95. If we look at Figure 3.6, and again focus on nonblack males, we see that individuals in bad health are not that far off from their objective probability, while individuals in good or excellent health are severely overestimating their survival probability and driving up the average.

To summarize, we want to stress two observations: first, subjective beliefs are informative and correlated with objective probabilities. Second, subjective beliefs are biased. Subjective probabilities overestimate the health/survival gradient, with individuals in bad health underestimating their survival probability relative to individuals in good health. Hence, there is both a systematic error along the age gradient and, within cohort, along the health gradient.

Estimation of the subjective life expectancy process

In this section, we use the health transitions and survival probabilities estimated in Foltyn and Olsson (2019) as a basis for estimating a different set of survival parameters that govern *subjective* survival beliefs. We take as given the parameters controlling health-to-health transitions conditional on survival since the HRS does not elicit any beliefs about future health states.

As explained above, the underlying data for this exercise takes the following form: HRS respondents are asked at date t to state their beliefs about surviving to a certain target age $\bar{a}$ (for example 75 or 85), which we reinterpret as the probability of being alive in period T, with $T = t + (\bar{a} - a_t)$. Thus, an observation i is given by the tuple

$$(p_i, h_i, \boldsymbol{x}_i, T_i)$$

where h_i denotes current health state, $\boldsymbol{x}_i$ is a vector of covariates including age and p_i is the subjective survival belief. We treat multiple observations from one individual independently: say a respondent ℓ is surveyed on survival beliefs to horizons T^1 and

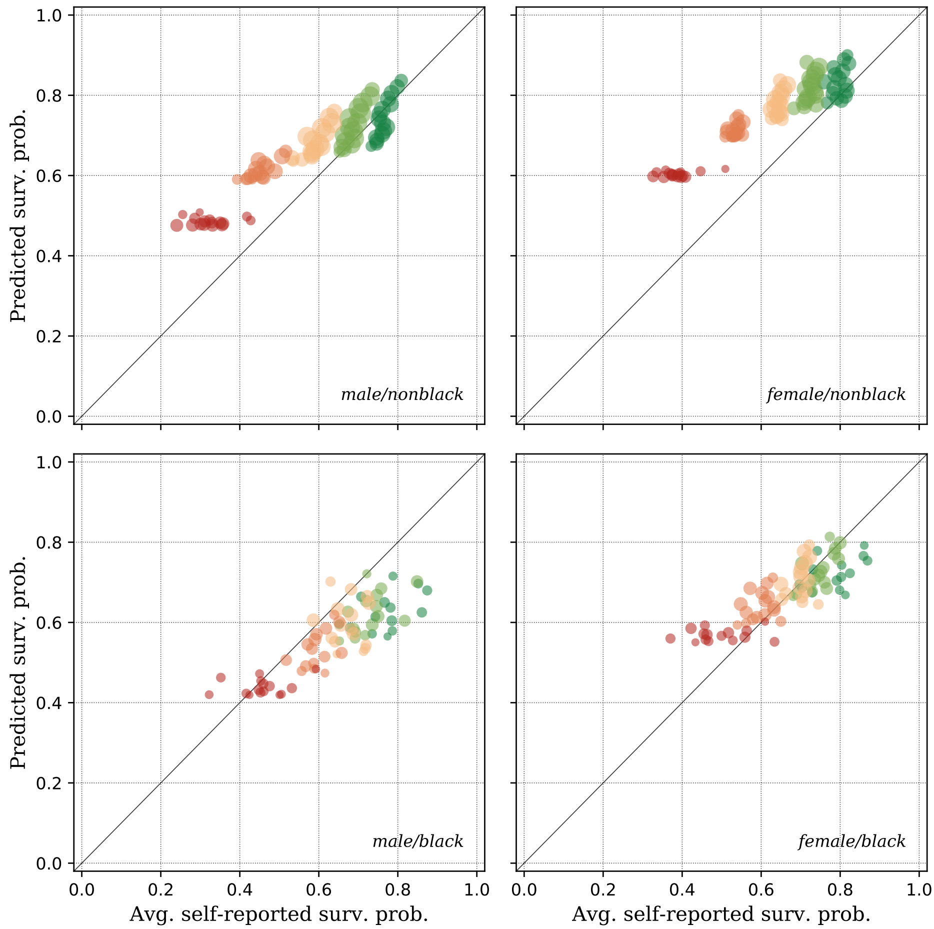


Figure 3.4: Objective vs. subjective survival probabilities for *target age 75*. Each dot represents the average for a gender/race/age/health group. The x-axis shows the average self-reported survival probability for that group, the y-axis the predicted (objective) survival probability. The color of the dot indicates health status, with red being poor health and green being excellent health.

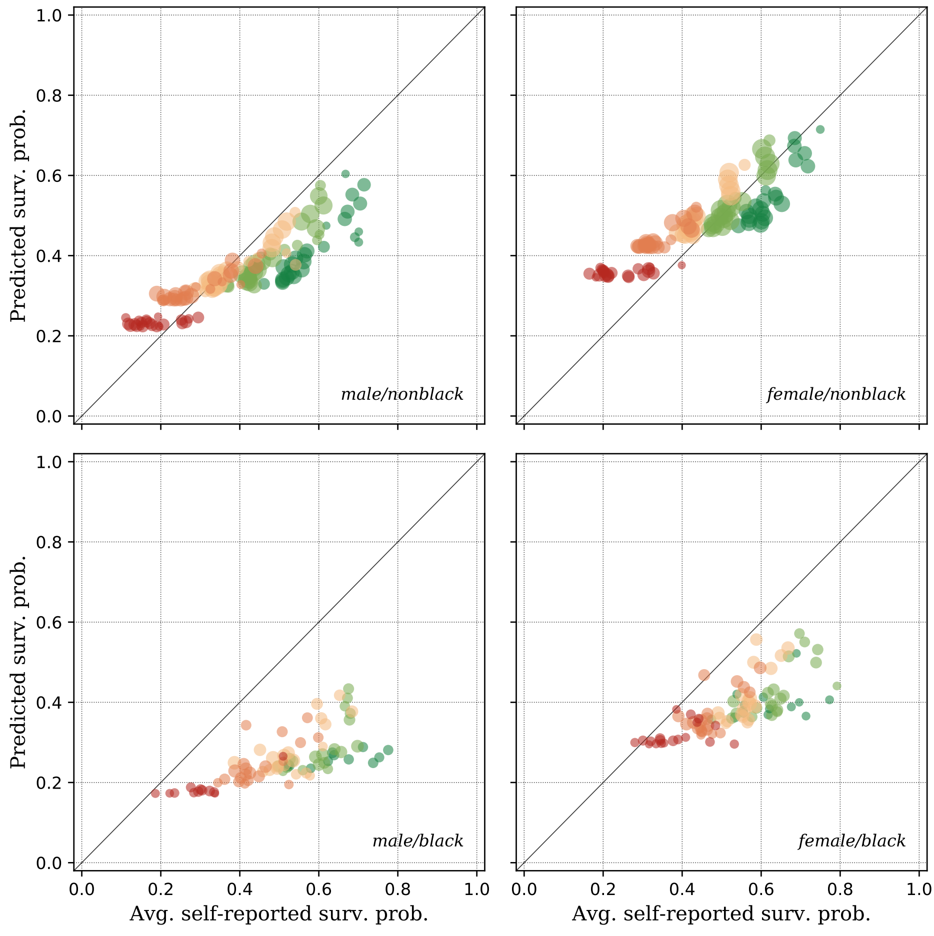


Figure 3.5: Objective vs. subjective survival probabilities for *target age 85*. Each dot represents the average for a gender/race/age/health group. The x-axis shows the average self-reported survival probability for that group, the y-axis the predicted (objective) survival probability. The color of the dot indicates health status, with red being poor health and green being excellent health.

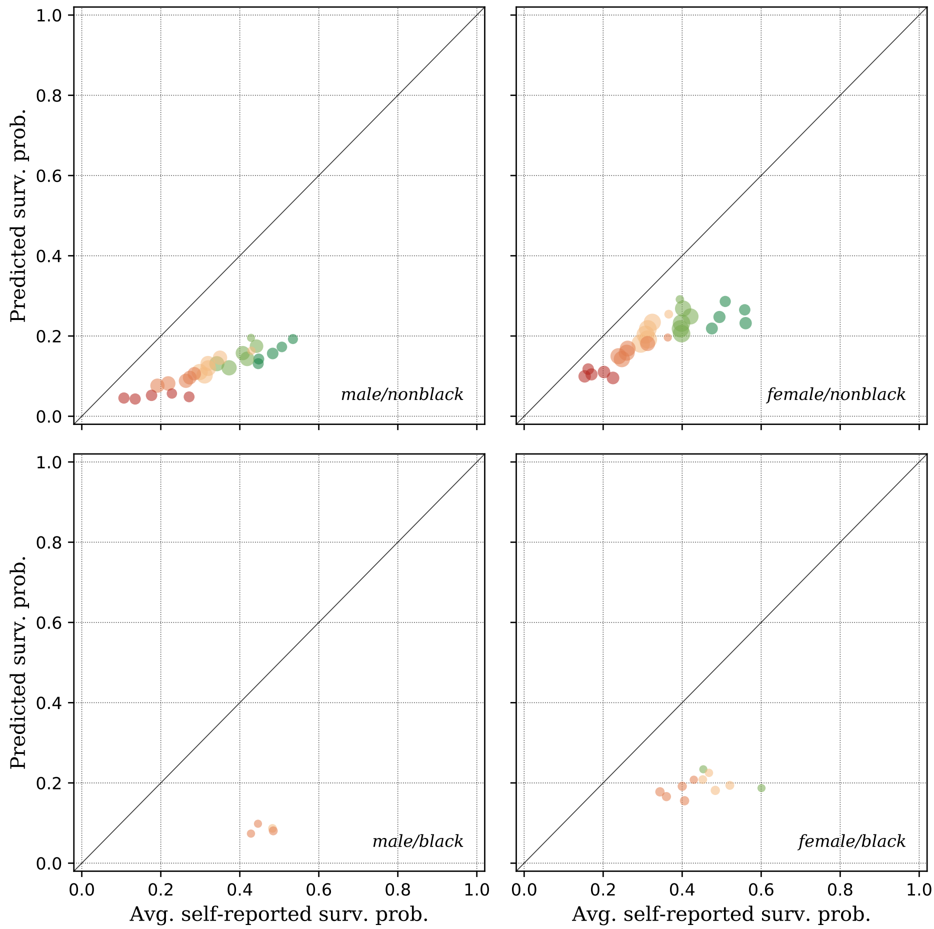


Figure 3.6: Objective vs. subjective survival probabilities for *target age 95*. Each dot represents the average for a gender/race/age/health group. The x-axis shows the average self-reported survival probability for that group, the y-axis the predicted (objective) survival probability. The color of the dot indicates health status, with red being poor health and green being excellent health.

T^2 in calendar years t_1 and t_2. This gives rise to the data

$$\left(p_{\ell t_1}, h_{\ell t_1}, \boldsymbol{x}_{\ell t_1}, T^1_{\ell t_1}\right)$$
$$\left(p_{\ell t_1}, h_{\ell t_1}, \boldsymbol{x}_{\ell t_1}, T^2_{\ell t_1}\right)$$
$$\left(p_{\ell t_2}, h_{\ell t_2}, \boldsymbol{x}_{\ell t_2}, T^1_{\ell t_2}\right)$$
$$\left(p_{\ell t_2}, h_{\ell t_2}, \boldsymbol{x}_{\ell t_2}, T^2_{\ell t_2}\right)$$

which we treat as four independent observations (except when bootstrapping confidence intervals, which we cluster at the individual level).

We are interested in estimating a parameter vector $\boldsymbol{\gamma}$ that captures these subjective beliefs conditional on $(h, \boldsymbol{x}, T)$. Assume that the i-th individual forms T-year-ahead survival beliefs based on the model

$$p^s_{it} = \phi_T\Big(h_{it}, \boldsymbol{x}_{it}, \boldsymbol{z}_{it}\Big) \tag{3.2}$$

where ϕ_T is an unknown nonlinear function that maps $(h, \boldsymbol{x}, \boldsymbol{z})$ into $[0, 1]$. The respondent's belief is allowed to depend on a vector of additional covariates $\boldsymbol{z}$ that are either unobserved or not included in our postulated model of survival.

In what follows we partition the sample into groups indexed by g, such that each unique combination of (h, x, T) forms a separate group. Denote by Γ_g all individual/year observations that satisfy

$$\Gamma_g = \Big\{(i, t) \,\Big|\, h_{it} = h_g, \boldsymbol{x}_{it} = \boldsymbol{x}_g, T_{it} = T_g\Big\}$$

i.e.,, all observations where the individuals are of the same age and have the same covariates, are in the same health state, and state their beliefs about survival to the same target age. Denote by $\overline{p}^s_g$ the (weighted) sample average of reported survival beliefs conditional on $(h_g, \boldsymbol{x}_g, T_g)$, i.e.,

$$\overline{p}^s_g = \frac{\sum_{(i,t)\in\Gamma_g} w_{it} \times \phi_T\left(h_g, \boldsymbol{x}_g, \boldsymbol{z}_{it}\right)}{\sum_{(i,t)\in\Gamma_g} w_{it}} \tag{3.3}$$

with w_{it} denoting sampling weights.

Now consider the nested-logit model counterpart of (3.3), which we denote by

$$\widehat{p}^s_g = \Pr\Big(s_{T_g} = 1 \,\Big|\, h = h_g, \boldsymbol{x}_g, T_g, \boldsymbol{\gamma} \Big)$$

i.e., the predicted probability of being alive for group g which is parametrized by the vector $\boldsymbol{\gamma}$. The observed sample moment for each group can then be written as

$$\overline{p}^s_g = \widehat{p}^s_g + u_g$$

where u_g is the residual that is not explained by our model. At this point, we are not imposing any restrictions on the residual. In particular, we are not postulating that this is an average of group-g individuals' forecast errors, since the individual forms beliefs according to (3.2). Our aim is to minimize these group-specific residuals using the least-squares objective function

$$J(\gamma) = \frac{1}{W} \sum_{g=1}^{N_G} W_g \left(\overline{p}_g^s - \widehat{p}^s \left(h_g, \boldsymbol{x}_g, T_g \,\middle|\, \gamma \right) \right)^2 \tag{3.4}$$

where $W_g = \sum_{(i,t) \in \Gamma_g} w_{it}$ is the sum of weights in group g. The estimated vector $\widehat{\gamma}$ is hence the arg min of $J(\gamma)$.

The resulting survival probabilities for nonblack men are shown in Figure 3.7 on the right, juxtaposing the objective survival probabilities estimated in Foltyn and Olsson (2019) on the left.[6] As can be seen, the subjective belief about survival while in health state "excellent" or "very good" is 100%. This does not mean that individuals in one of those health states believe that they will live forever, but rather that death is preceded by a deterioration in health.

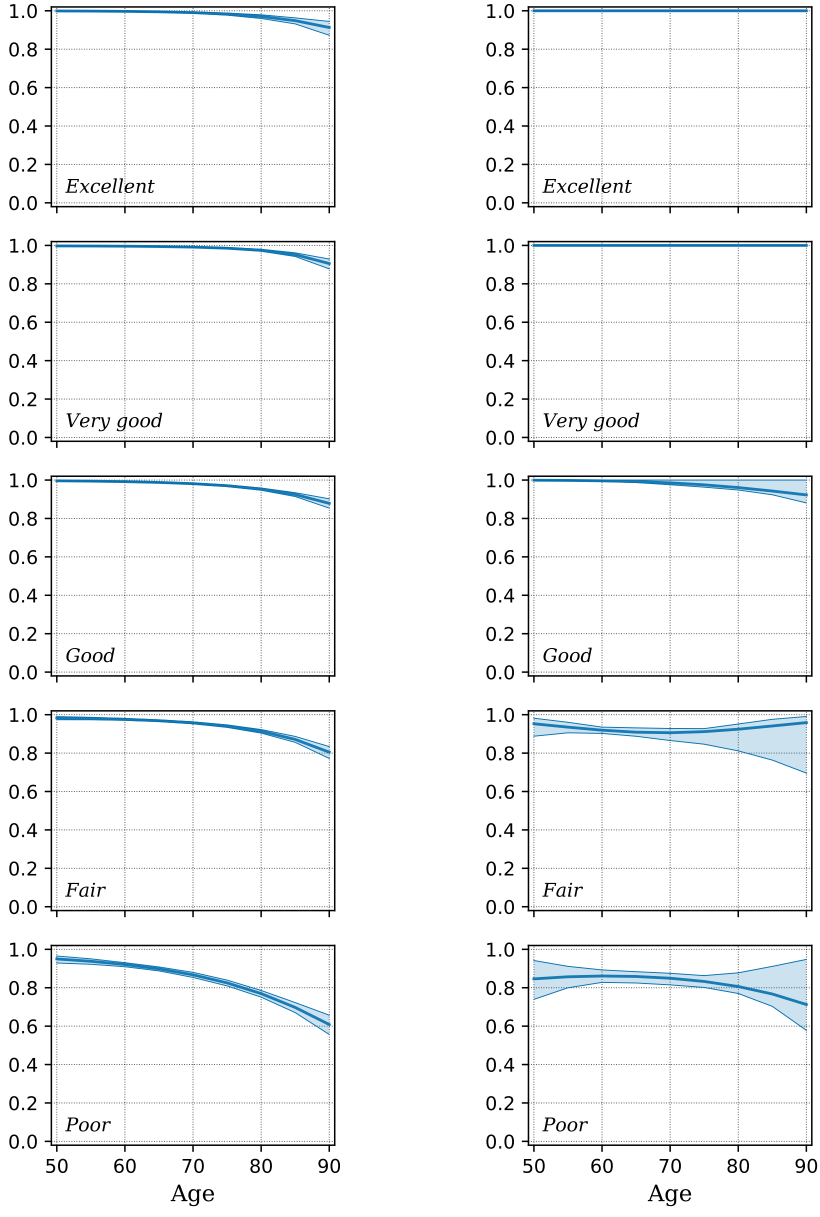
Note that ex ante, it is not obvious that a dynamically consistent belief system exists that can match the elicited subjective beliefs.[7] It is easy to imagine subjective beliefs that are contradictory: for instance, if 50-year-olds say that the probability of surviving until the age of 75 is 80%, and the probability of surviving until the age of 85 is 40%, this implies that the probability of surviving until 85, *conditional on turning 75*, is 50%. If then the 75-year-olds said that their probability of surviving until the age of 85 is 80%, we would have a belief system that is dynamically inconsistent.

However, it turns out that the subjective beliefs can actually be mapped into a dynamically consistent set of survival beliefs conditional on age and health. In Figure 3.8 we plot the model-predicted subjective survival against elicited beliefs. As can be seen, the estimated model for subjective beliefs captures the main picture, since the dots, each representing an age/health/target-age group, line up reasonably close to the 45 degree line.

Figure 3.9 summarizes the results, showing the life expectancy by age and health state using the objective and the subjective survival process. As can be seen, at all ages, the difference in life expectancy between the best and the worst health state is larger when using subjective life expectancies. The divergence between objective and

6. Results for other demographic groups are available upon request. For the model in the next section, we will use the estimates for nonblack men.

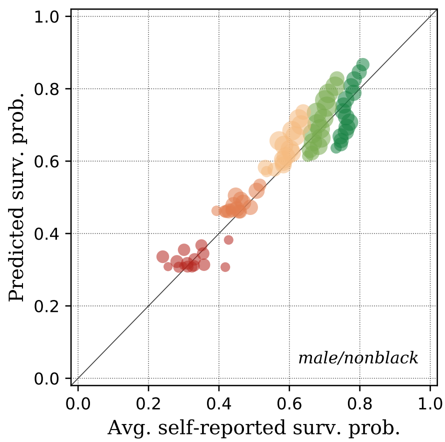
7. As an alternative to dynamically consistent beliefs, one could for instance model survival beliefs as the result from a Bayesian learning process, as in Groneck, Ludwig, and Zimper (2016), who use a model of Choquet Bayesian learning of survival beliefs, allowing for likelihood insensitivity. However, since our estimated subjective belief system matches the elicited beliefs well, we think this is appropriate for the purposes of this paper.



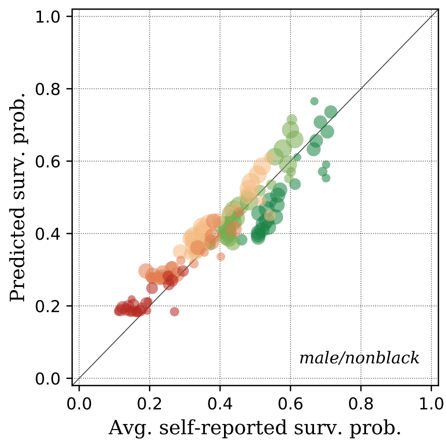
(a) Objective probabilities

(b) Subjective probabilities

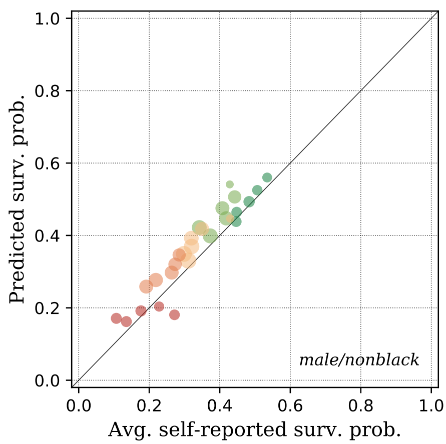
Figure 3.7: One-year survival probability (model estimates). Shaded areas indicate bootstrapped 95% confidence intervals. For each bootstrapped sample we re-estimate the objective health process.



(a) Target age 75

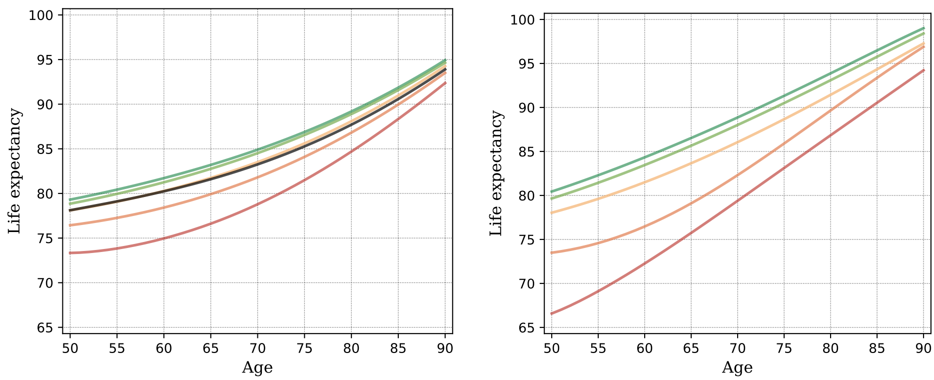


(b) Target age 85



(c) Target age 95

Figure 3.8: Elicited beliefs about survival vs. estimated *subjective* survival probabilities. Each dot represents the average for an age/health group. The x-axis shows the average self-reported survival probability for that group and the y-axis the predicted survival probability according to the subjective model. The color of each dot indicates health status, with red being poor health and green being excellent health.



(a) Objective probabilities

(b) Subjective probabilities

Figure 3.9: Life expectancy by age and health state. The color indicates health state: green is excellent while red is poor health.In the left graph, showing the life expectancy using objective probabilities, the black line indicates the average in the population, weighted by the distribution over health states.

subjective life expectancies is particularly large for individuals in bad health states, who substantially underestimate survival at younger ages. Conversely, individuals in all health states overestimate their chances of survival late in life.

3.3 Model

In this section, we describe the overlapping generations model we use for our analysis. Time is discrete and every time period is assumed to be one year. Agents derive utility from consumption and face three types of idiosyncratic risks: shocks to persistent productivity, transitory productivity shocks and shocks to health and survival. Agents can only save in a riskless bond, and they face an exogenous borrowing constraint.

The agent's problem

There is a unit mass of individuals distributed across N_t cohorts according to the ergodic distribution implied by the transition matrix of survival probabilities. An individual of age $t \in \{1, \ldots, N_t\}$ and health $h \in \{1, \ldots, N_h\}$ has a one-period survival probability to age $t+1$ given by π^s_{th}, with $\pi^s_{th} = 0$ for $t = N_t$ regardless of health state.

Individuals are assumed to be in the labor force for the first $T_R - 1$ years of their life and exogenously retire in the period when they attain age T_R. While in the labor force, they are hit by persistent and transitory labor productivity shocks. During

retirement, individuals receive social-security retirement benefits which depend on their last persistent labor state in working age.

Bequests are distributed to new-born individuals in a dynastic way: one young individual receives the bequests from one dying individual.

Retired agents. A retired individual with idiosyncratic states (a, p, h, t), where a is cash-at-hand, p is the (fixed) persistent component of labor earnings, h is the current health state, and t is the age, maximizes utility according to

$$V_R(a, p, h, t) = \max_{c,\, b'} \left\{ u(c) + \beta \left(\pi^s_{th} \mathbf{E}\left[V_R(\boldsymbol{x}') \,\middle|\, h, t \right] + (1 - \pi^s_{th}) V_b(a'_b) \right) \right\} \quad (3.5)$$

subject to the constraints

$$\begin{aligned} a &\geq c + b', \quad c \geq 0, \quad b' \geq 0 \\ a' &= Rb' + \iota'_R \\ \iota'_R &= y'_R w - T_y(y'_R w) \\ y'_R &= \omega_{T_R - 1} p_R(p) \bar{\epsilon} \\ a'_b &= Rb' - T_b(Rb') \end{aligned}$$

where $\boldsymbol{x}' = (a', p, h', t + 1)$ is the continuation state conditional on survival. Next-period after-tax retirement income is denoted by ι'_R and depends on the non-linear tax schedule $T_y(\bullet)$. $p_R(\bullet)$ is a function mimicking the regressive replacement rate of the U.S. social security system, w is the economy-wide wage rate, $\bar{\epsilon}$ is the average of the transitory earnings shocks hitting the working-age population and ω_{T_R-1} is the value of the deterministic age profile of earnings just prior to retirement. Bequests are subject to a potentially non-linear tax schedule $T_b(\bullet)$. The gross return on savings is given by $R = 1 + r - \delta_k$.

Specifically, in the terminal period with $t = N_t$, the individual solves

$$\begin{aligned} V_R(a, p, h, t) &= \max_{c,\, b'} \left\{ u(c) + \beta V_b(a'_b) \right\} \\ \text{s.t.} \quad a &\geq c + b', \quad c \geq 0, \quad b' \geq 0 \\ a'_b &= Rb' - T_b(Rb') \end{aligned} \quad (3.6)$$

Working-age households. Non-retired individuals are assumed to be in the labor force, so we use the subscript LF to denote their value and policy functions. The vector of idiosyncratic state variables is $\boldsymbol{x} = (a, p, h, t)$.

A working-age individual draws a persistent and a transitory labor shock component that together with a deterministic age profile of earnings, ω_t, determine his

labor productivity in the given period. The persistent component p takes on the values from the set $\mathcal{P}$, while the transitory shock realizations ϵ are drawn from $\mathcal{E}$.

A working-age individual with $t < T_R - 1$ who will continue to be in the labor force next period solves

$$V_{LF}(\boldsymbol{x}) = \max_{c,\, b'} \left\{ u(c) + \beta \left(\pi^s_{th} \mathbf{E}\left[V_{LF}(\boldsymbol{x}') \,\middle|\, p, h, t \right] + (1 - \pi^s_{th}) V_b\left(a'_b\right) \right) \right\}$$

$$\begin{aligned}
\text{s.t.} \quad a &\geq c + b', \quad c \geq 0, \quad b' \geq 0 \\
a'_b &= Rb' - T_b\left(Rb'\right) \\
a' &= Rb' + \iota' \\
\iota' &= \left[y' - T_{ss}\left(y'\right)\right]w - T_y\left(\left[y' - T_{ss}\left(y'\right)\right]w\right) \\
y' &= \omega_{t+1} p' \epsilon'
\end{aligned}$$

where $\boldsymbol{x}' = (a', p', h', t+1)$. We denote the earnings of working individuals, net of income taxes $T_y(\bullet)$ and payroll taxes $T_{ss}(\bullet)$, as ι'.

In the final period of their working life, i.e., when $t = T_R - 1$, individuals in the labor force solve

$$V_{LF}(a, p, h, t) = \max_{c,\, b'} \left\{ u(c) + \beta \left(\pi^s_{th} \mathbf{E}\left[V_R\left(a', p, h', t+1\right) \,\middle|\, h, t \right] + (1 - \pi^s_{th}) V_b\left(a'_b\right) \right) \right\}$$

subject to

$$\begin{aligned}
a &\geq c + b', \quad c \geq 0, \quad b' \geq 0 \\
a'_b &= Rb' - T_b\left(Rb'\right) \\
a' &= Rb' + \iota'_R \\
\iota'_R &= y'_R w - T_y\left(y'_R w\right) \\
y'_R &= \omega_{T_R - 1} p_R(p) \bar{\epsilon}
\end{aligned}$$

which is identical to the retired individual's problem.

Technology

The production side of the model is standard. Competitive firms employ labor and capital hired from households to produce a homogeneous final good, which is used for both consumption and investment. The aggregate production function is assumed to be Cobb-Douglas:

$$F(K, L) = K^{\alpha_k} L^{1-\alpha_k}.$$

Capital depreciates at the rate δ_k.

Government

We assume that the government runs a PAYGO social security system that has to balance in each period, and that remaining (wasteful) government expenditures have to be fully financed by estate and income taxes. We first describe the social security system and thereafter the general government budget.

Social security system

We use a stylized version of the actual retirement income formula used in the U.S. social security system. It captures the main features, such as a regressive replacement rate based on pre-retirement income and a cap for maximum benefits. In the model, we define retirement benefits to be a product of the economy-wide wage rate, the life-cycle profile wage component from the last year before retiring, the average transitory component, and a function that mimics the regressive replacement rate of the U.S. social security system:

$$\iota_R(p) = w \times y_R(p) = w \times \omega_{T_R-1}\bar{\epsilon} p_R(p)$$

where the replacement function $p_R(\cdot)$ is given by

$$p_R(p) = \begin{cases} \rho_1 p & \text{if } p \leq \widetilde{p}_1 \\ \rho_1 \widetilde{p}_1 + \rho_2 (p - \widetilde{p}_1) & \text{if } \widetilde{p}_1 < p \leq \widetilde{p}_2 \\ \rho_1 \widetilde{p}_1 + \rho_2 (\widetilde{p}_2 - \widetilde{p}_1) + \rho_3 \left(\min\left\{ \widetilde{p}_{max}, p \right\} - \widetilde{p}_2 \right) & \text{else} \end{cases}$$

where $\widetilde{p}_1$ and $\widetilde{p}_2$ are bend points and $\widetilde{p}_{max}$ the contribution and benefit base (CBB) in the social security income formula, expressed in terms of the individual's permanent labor state. In the appendix, section 3.B, the transformation between actual replacement rate bend points and CBB expressed in USD ($b_1^\$$, $b_2^\$$, and $e_{max}^\$$) and the model counterparts are described in detail, as well as the derivation of total government expenditures on retirement, G_{ss}.

The government expenditures on retirement are financed by a payroll tax. The payroll tax function is defined as

$$T_{ss}(y) = \tau_{ss} \times \min\{y_{max}, y\}$$

where y_{max} expresses maximum taxable earnings in terms of labor productivity, i.e.,

$$y_{max} = \left(e_{max}^\$ / e_{med}^\$ \right) y_{med}$$

where $e_{max}^\$$ is the contribution and benefit base expressed in USD as above, and $e_{med}^\$$ are the median earnings in the reference year.

The derivation for total payroll taxes raised in each period, $\overline{T}_{ss}$, can be found in the appendix, section 3.B. To balance the social security system, we need to find τ_{ss} such that $G_{ss} = \overline{T}_{ss}$.

Government budget

We assume that the government raises estate taxes and labor income taxes to finance non-discretionary expenditures that amount to a constant fraction g of output. In the following section, we will first describe the estate tax, thereafter the income tax, and finally how the government budget is balanced.

Estate taxes. The estate tax schedule is defined as

$$T_b(b) = \begin{cases} 0 & \text{if } b \leq \chi_b \\ \frac{\tau_b}{2}\left[\sin\left(\pi\left[\frac{b-\chi_b}{B} - 1\right]\right)\frac{B}{\pi} + b - \chi_b\right] & \text{if } \chi_b < b \leq \chi_b + B \\ \tau_b(b - \chi_b - B) + \frac{\tau_b}{2}B & \text{else} \end{cases}$$

This formulation is effectively a step function so that estates valued at less than χ_b are exempt from taxes. For values b in an interval $\chi_b < b \leq \chi_b + B$, the marginal tax rate is increasing, and for $b > \chi_b + B$, the marginal tax rate is τ_b. The tax schedule is twice continuously differentiable, which is required by our solution algorithm. Details about the estate tax are given in the appendix, section 3.C.

Income taxes. For income taxes, we adopt the same tax function as in Heathcote, Storesletten, and Violante (2017), which is defined as

$$T_y(\iota) = \iota - \lambda\iota^{1-\tau} \tag{3.7}$$

where ι is either earnings (net of payroll taxes) or retirement income. We assume that the progressivity parameter τ is fixed, and we pin down λ such that the government budget is balanced in each period.

Total income taxes raised by the government amount to the sum of income taxes raised from the employed and from the retired individuals:

$$T_{inc} = T_e + T_R$$

The derivation of the total income tax raised by the government in each period can be found in the appendix, section 3.C.

Government budget balance. The non-discretionary expenditures amount to a constant fraction g of output: $G = gY$. The government budget is balanced by solving for a λ in (3.7) such that $G = T_{inc} + T_b$ holds.

Equilibrium definition

A recursive competitive equilibrium is given by a set of prices $\{R, w\}$, tax rates $\{\tau_{ss}, \lambda\}$, decision rules $C(a, p, h, t)$ (for consumption) and $B(a, p, h, t)$ (for savings), and a stationary distribution Γ such that:

1. The decision rules solve the agents' problem for all (a, p, h, t).

2. Factor prices are given by:

$$r = F_1(K, L) \qquad \text{and} \qquad w = F_2(K, L)$$

3. τ_{ss} and λ are set so that both the social security system and the general government budget balance.

4. Capital and labor markets clear:

$$K' = \int B(a, p, h, t)\, \mathrm{d}\Gamma \qquad \text{and} \qquad L = \sum_{t=1}^{T_R-1} \sum_{p} \sum_{\epsilon} \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon)\omega_t p\epsilon$$

where $\mu_t(t)$, $\mu_p(p)$ and $\mu_\epsilon(\epsilon)$ are the ergodic distributions over age, the persistent and the transitory labor shocks, respectively.

5. The distribution Γ is stationary, i.e., for all relevant Borel sets $\mathcal{B}$

$$\Gamma(\mathcal{B}, p, h, t) = \sum_{\overline{p}} \sum_{\overline{h}} \sum_{\overline{t}} \pi(p|\overline{p})\pi(h|\overline{h})\pi(t|\overline{t}) \int_{a:B(a,\overline{p},\overline{h},\overline{t})\in\mathcal{B}} \Gamma(da, \overline{p}, \overline{h}, \overline{t})$$

3.4 Calibration

In this section, we explain our strategy to calibrate the model. First, we parametrize the instantaneous utility function and the bequest motive. Then, we adopt a threefold strategy. First, some parameters are taken directly from the literature. Second, parameters guiding the health and death process are directly estimated from the data as reported in section 3.2. Lastly, a set of parameters are estimated with the method of simulated moments.

Preferences

We assume that utility in each period is of the standard CRRA form with a risk-aversion parameter σ set to one.

In the macro literature, it is common to follow De Nardi 2004 and describe the bequest motive as

$$V_b(a) = \theta_B \frac{(a+\kappa)^{1-\sigma} - 1}{1-\sigma} \tag{3.8}$$

where θ_B determines the strength of the bequest motive, and κ determines to what extent bequests are a luxury good. In the model environment where bequests are not operative we let $\theta_B = 0$. For the alternative scenario with a warm-glow bequest motive, we choose parameter values for θ_B and κ to match the life-cycle asset profiles of the elderly as observed in the HRS.

Externally calibrated parameters

Demographics and the life cycle. Agents are assumed to enter the economy at age 20, and retire at the age of 65, hence $T_R = 46$. The maximum age an agent can reach is 99, hence $N_t = 80$.

Age-dependent wage profile and idiosyncratic earnings risk. We assume that the log-labor earnings of an individual in the labor force follow a process with transitory and persistent shocks:

$$\log y_t = \log \omega_t + \log p_t + \log \epsilon_t$$

where ω_t is the age profile part, p_t is the persistent component and ϵ_t is the transitory component of earnings. The persistent component is assumed to follow an AR(1) process specified as:

$$\log p_t = \rho \log p_{t-1} + \eta_t$$

with persistence ρ and innovation $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$. The transitory shock is given by $\log \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

Hence, the stochastic part of the wage process is characterized by the parameters $(\rho, \sigma_\eta^2, \sigma_\epsilon^2)$ which we set to $(0.9695, 0.0384, 0.0522)$, following Krueger, Mitman, and Perri 2016.[8]

We use the Rouwenhorst procedure to discretize the persistent part of the process into an eleven-state Markov chain, and we discretize the transitory shock into three states.

We choose the deterministic age profile of earnings estimated for high-school graduates by Cocco, Gomes, and Maenhout 2005, and renormalize it such that the average labor productivity is unity.

8. Note that Krueger, Mitman, and Perri 2016 remove the age effect before estimating this process and hence, we can use this stochastic process on top of the age-dependent profile.

Parameter	Description	Value	Source
Production technology parameters			
α_k	Capital share	36%	Standard value
δ_k	Depreciation rate	9.6%	Standard value
Social security			
ρ_1	Replacement rate bracket 1	90%	2013 SS rules
ρ_2	Replacement rate bracket 2	32%	2013 SS rules
ρ_3	Replacement rate bracket 3	15%	2013 SS rules
$b_1^{\$}$	Bendpoint 1 (in USD)	9,492	2013 SS rules
$b_2^{\$}$	Bendpoint 2 (in USD)	57,216	2013 SS rules
$e_{max}^{\$}$	CBB (in USD)	113,700	2013 SS rules
Government budget			
g	Gov. spending (share of GDP)	6%	Brinca et al. 2016
τ	Tax progressivity	0.137	Brinca et al. 2016
τ_b	Marginal tax on estates	30%	Authors' approximation[9]

Table 3.1: Calibrated parameters

Remaining externally calibrated parameters. The remaining parameters that are set externally are listed in Table 3.1. The bend points and the contribution and benefit base are reported in U.S. dollars, to facilitate the interpretation. Details for transforming them into values used in the model are found in the appendix, section 3.B.

Health and death process

We use the process for health transitions and death probabilities described in section 3.2. However, agents enter the model at the age of 20, but the health and death process we estimated starts at the age of 50. Therefore, we make the following assumptions: everyone is born in health state "excellent". Thereafter, we use the health transition matrix for the age of 50 to roll forward the population, assuming certain survival. At the age of 50, agents start facing a positive probability of death according to our estimated process. The resulting cohort sizes and distribution of health states are shown in Figure 3.10.

Estimated parameters

Given the parameters and processes described above, we estimate the remaining parameters with the method of simulated moments, i.e., we minimize the weighted sum of squared distances between targeted and simulated moments. For this exer-

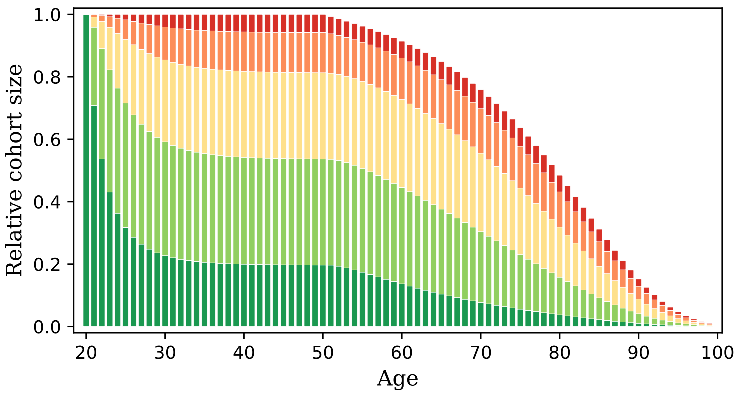


Figure 3.10: Cohort size and health state distribution. The color indicates health state: green is excellent while red is poor health.

cise, we use the benchmark model where we let all agents have the same survival expectations.

Target moments. In the first model, with no bequest, the capital-to-output ratio is the only target moment used. We target a level of 3.0, by choosing the appropriate value of β, the discount factor. In the model with a bequest motive, we still target a capital-to-output ratio of 3.0, but we also match the life-cycle profile of assets. We use the median wealth levels at ages 55, 60, 65, 70, 75, 80 and 85. The capital-to-output ratio and the life-cycle profile are jointly matched by choosing the parameters for the discount factor (β), the bequest utility weight (θ_B), the bequest utility shifter (κ), and the estate tax exemption (χ_b).

Estimation results. The resulting parameter estimates are shown in Table 3.2. In the estimation of the model without bequest, the resulting capital-to-output ratio is 3.0. In the estimation of the model with bequest, the resulting capital-to-output ratio is also 3.0, and the resulting asset holdings by age and their data counterparts are shown in Table 3.3. As can be seen, the resulting wealth level in the model for the age of 65, i.e., exactly when the model agents enter retirement, is slightly higher than what is observed in the data. The reason is the fixed retirement age. Since all agents retire at the exact same age, they will all accumulate assets until exactly that point, while in reality, people retire more gradually.

To assess the bequest motive, it is informative to look at non-targeted moments. The bequest-to-wealth ratio in the model is 1.9%. It is difficult to precisely measure

Parameter	Description	Value
Model without bequest		
β	Discount factor	0.981
Model with bequest		
β	Discount factor	0.961
θ_B	Bequest weight	10.33
κ	Bequest shifter	1.98
χ_b	Estate tax exemption	27.32

Table 3.2: Estimated parameters

Median asset holdings							
Age	**55**	**60**	**65**	**70**	**75**	**80**	**85**
Data	1.30	1.65	1.69	1.73	1.62	1.47	1.41
Model	1.26	1.61	2.13	1.78	1.61	1.43	1.35

Table 3.3: Life-cycle asset profile in the model with a bequest motive, relative to the median asset holdings at age 50.

this figure in the data, but according to Gale and Scholz 1994, using SCF data, it should be closer to 0.9%. However, according to others, 2% is "a conservative estimate".[10] Another measure is the fraction of deceased individuals who pay estate taxes. The model result is 0.74%. This figure has varied during the last 15 years: in 2004 it was 0.8% and in 2013 0.2%.[11] Hence, the predictions from our model regarding these two measures seem to be reasonably in line with the data.

A last measure is estate tax revenue as a fraction of GDP. In the model, this figure is 0.02%. In the data, this figure has varied during the last few years, from a high of 0.17% in 2007 to a low of 0.07% in 2011.[12] Hence, for this value, our model prediction is at the lower end; however, this figure is to a large extent a result of the upper tail of the asset distribution, something that we know that our model does not fully capture.

10. See Larry Summer's opinion at https://www.reuters.com/article/column-summers/column-how-to-target-untaxed-wealth-lawrence-summers-idUSL1E8NG2MC20121216?irpc=932.

11. Calculated as the ratio of the number of estates subject to estate taxes, as reported by the IRS, and the number of deaths taken from CDC records.

12. Calculated as the ratio of net estate taxes paid, as reported by the IRS, and official GDP figures.

3.5 Results

As mentioned above, we solve the model under three distinct assumptions about the survival beliefs:

1. *No survival heterogeneity (NSH).* All individuals face the same average survival expectations.
2. *Objective survival heterogeneity (OSH).* Individuals have heterogeneous objective survival expectations.
3. *Subjective survival heterogeneity (SSH).* Individuals have heterogeneous subjective survival expectations.

The first scenarios serve as our baseline: this is the standard model where all agents face the same survival risk and hence, all agents have the same effective discount factor, conditional on age. For this scenario, we collapse the health states into only one, and use the average survival rates.

In the second scenario, we use the objective process for health transitions and survival probabilities described in section 3.2. Hence, in this model, people are perfectly informed about their true survival probability conditional on health and age.

In the third scenario, agents believe and act according to the subjective survival process we estimated in section 3.2. However, this subjective process does not correspond to the true survival process, which is what we use when simulating the model.

The section is structured as follows. We will first analyze a model without a bequest motive. We discuss how the solution to the agent's problem changes between the three scenarios (with no survival heterogeneity, objective survival heterogeneity, and subjective survival heterogeneity), and then aggregate the changed behavior by the individuals into economy-wide general equilibrium effects.

Thereafter, we repeat the analysis, but this time using a model with a bequest motive, and see how that affects the results.

Model without bequest motive

We start with a model in which the agents have no bequest motive and hence only save for their own consumption needs. The main effect from introducing health and survival heterogeneity is that agents in bad health save less than their healthy counterparts.

The reason why individuals in bad health save less is straightforward: they expect to live a shorter life. The same but opposite effect is present for individuals in good health: their life expectancy is longer, and therefore they save more. This is true regardless of whether agents compute their life expectancy using objective survival probabilities or subjective beliefs. However, the difference in the savings rate between

agents in bad and good health increases if we let the agents act according to their subjective beliefs. The reason is the bias we documented in section 3.2: individuals in bad health underestimate their survival probability more (or, for some ages, do not overestimate to the same extent) as compared to agents in good health.

These patterns are documented in Figure 3.11, which shows the difference in the savings rate for the baseline model without any bequest motive. The bars show, by health state, the difference in the savings rate (in percentage points) for the model with objective heterogeneity (in darker color) compared to the baseline model with no heterogeneity in survival. The lighter color shows the additional effect of adding subjective beliefs. The horizontal axis represents the cash-at-hand percentiles for the age in question in the baseline model with no heterogeneity.[13]

As can be seen, for 50-year-olds, the difference is most pronounced for individuals in bad health, as indicated by the red bars. With objective survival heterogeneity, individuals in the 10th cash-at-hand percentile in poor health have a savings rate that is 9 percentage points lower than individuals of the same wealth in the model with no heterogeneity. Adding subjective beliefs, the difference is magnified: the savings rate of an individual in the 10th percentile is now 22 percentage points lower than for an individual with the same wealth with average survival expectations.

The reason for the bigger impact of the subjective belief on the individuals in worst health can be found by looking at Figure 3.7. At lower ages, the subjective and the objective survival expectations are very similar for individuals in the better health states and therefore the additional effect of adding subjective beliefs is small. However, individuals in worse health are severely underestimating their longevity and thus the additional effect on the savings rate is larger.

The same pattern can be seen in the graphs for 60-year-olds and 70-year-olds, even though for the latter, the additional effect from adding the subjective beliefs is slightly weaker if we continue to focus on the individuals in poor health. The reason for this can be found in Figure 3.5: at the age of 70, individuals start taking into account the probabilities that are shown in this figure and, as can be seen, the subjective survival beliefs of the individuals in bad health are now closer to the objective probabilities.

However, at the age of 70, the effect of adding subjective beliefs starts showing up for individuals in good health; in other words, there is an additional effect of the subjective beliefs on the savings rate for healthy individuals, as indicated by the lighter green parts of the green bars. As can be seen from the graph of 80-year-olds, adding subjective beliefs at this age mainly affects the individuals in good health.

Figure 3.11 shows the change in the savings rate in percentage points, but this

13. The reason why we plot results by cash-at-hand percentiles (as opposed to the cash-at-hand level) is that the graphs then show the difference at the wealth levels that matter, i.e., where there is a positive mass of agents in equilibrium. A large difference in policy functions for very high wealth levels for the 90-year-olds is not that informative, since no 90-year-old will be that rich in equilibrium anyhow.

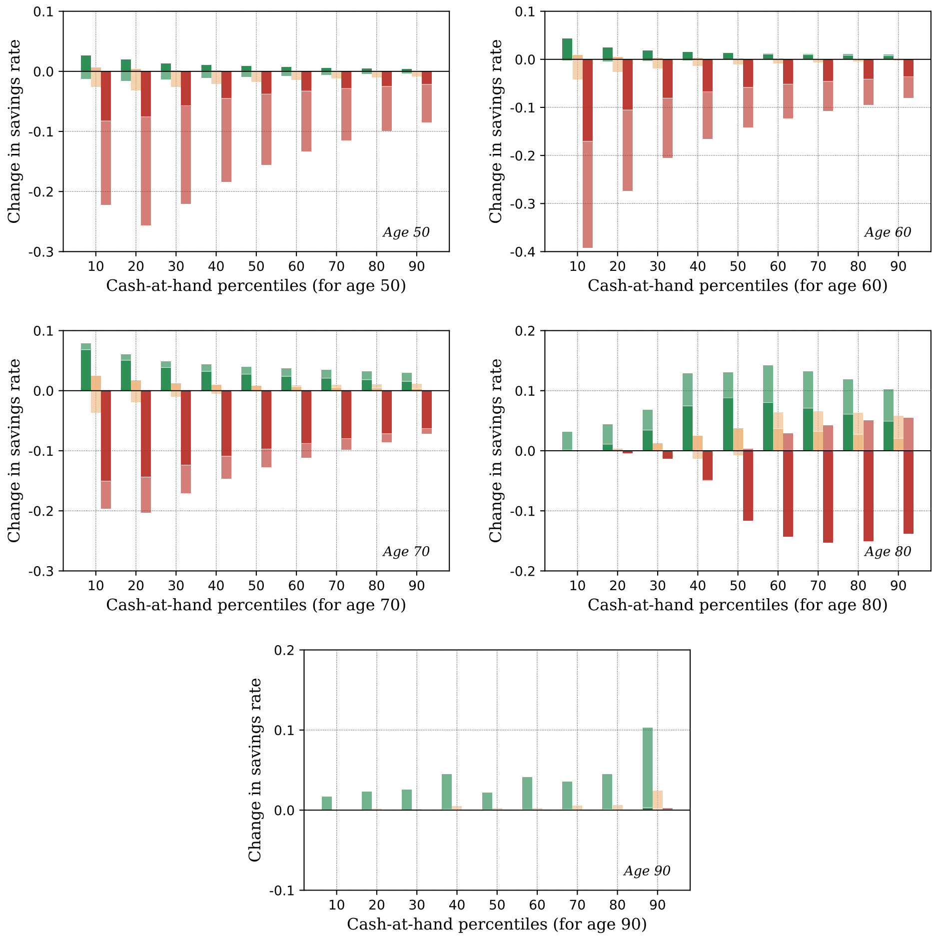


Figure 3.11: Difference in the savings rate for the model *without bequest motive* (percentage points/100, i.e.,, 0.1 indicates a difference of 10 pp). The darker color shows the model with objective survival heterogeneity compared to the baseline model with no survival heterogeneity. The lighter color shows the additional effect of adding subjective survival beliefs. The x-axis depicts cash-at-hand percentiles by age (equilibrium values for the baseline model). The color indicates health state: green is excellent while red is poor health.

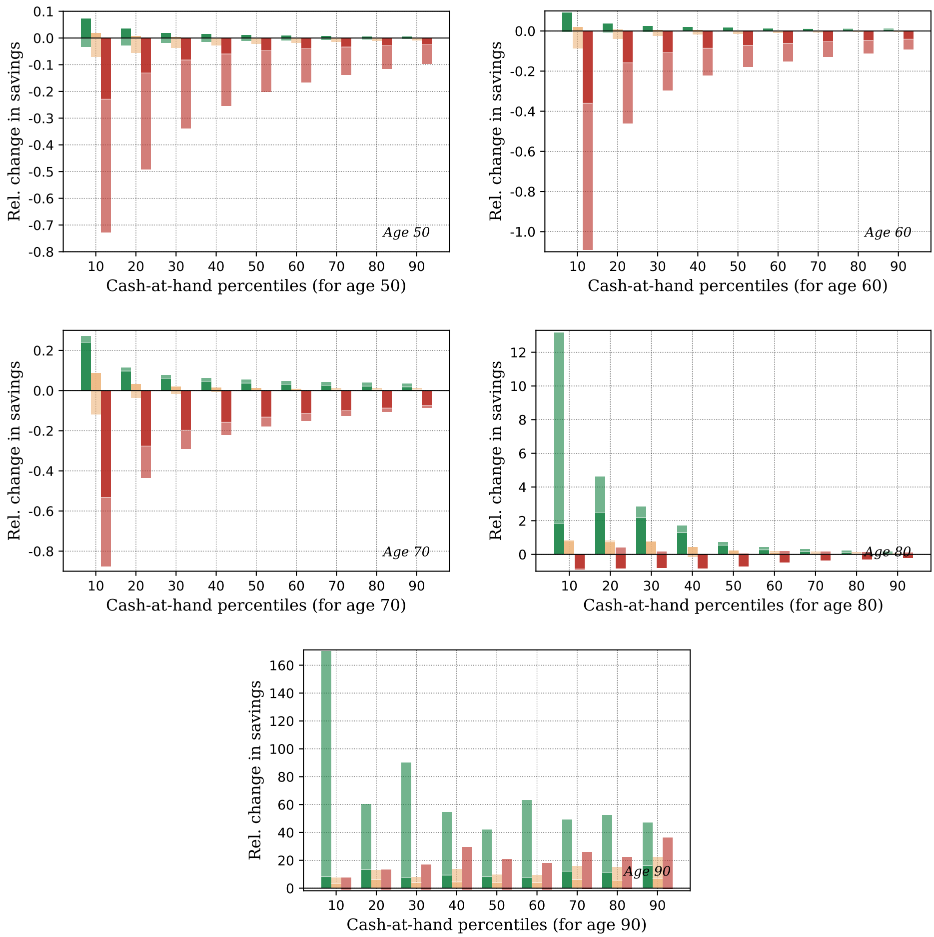


Figure 3.12: Relative change in savings for the model *without bequest motive* (percent/100, i.e.,, 0.1 indicates an increase by 10 percent). The darker color shows the model with objective survival heterogeneity compared to the baseline model with no survival heterogeneity. The lighter color shows the additional effect of adding subjective survival beliefs. The x-axis depicts cash-at-hand percentiles by age (equilibrium values for the baseline model). The color indicates health state: green is excellent while red is poor health.

might be misleading. For instance, at the age of 80, the savings rate is very low for the poorest individuals and hence even a substantial increase in relative terms looks like a small increase in percentage points. To better assess the magnitudes, we plot the relative changes in savings in Figure 3.12. The graph for 80-year-olds shows that the relative increase in savings for 80-year-olds in excellent health is large, even for the poorest individuals: it is an increase of 1200%. For the 90-year olds, the difference is enormous in relative terms.

Figure 3.13 shows the resulting life-cycle profiles for the three different scenarios: no survival heterogeneity (NSH), objective survival heterogeneity (OSH), and subjective survival heterogeneity (SSH). In all graphs, we integrate out the productivity dimension.

As expected, and as all three graphs show, the life cycle profile for wealth peaks at the age of 64, which is the last year before retirement. When the agents enter retirement, they start drawing down their wealth and the average individual who survives until the age of 90 has drawn down all of his savings (remember that all individuals receive retirement benefits, so they are not risking zero consumption).

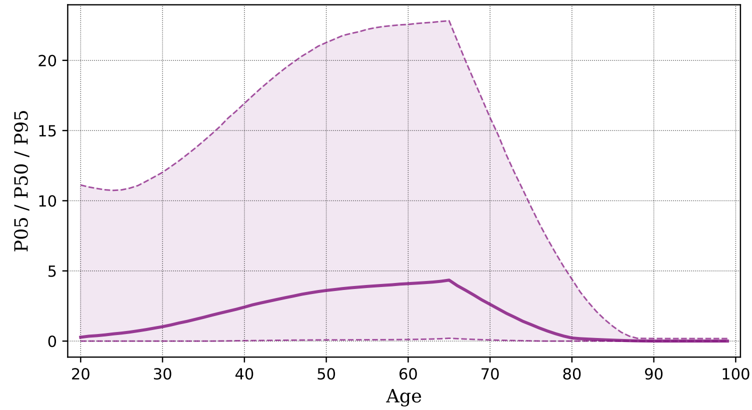
The profile for the OSH model illustrates that individuals in bad health have less wealth than individuals in good or excellent health. It should be noted that the mass of individuals in bad health is not static, but consists of individuals who have been in bad health for many periods, as well as individuals who just recently draw a bad health shock.

Moreover, the profile for the SSH model shows that the difference in wealth between individuals in good and bad health at the age of 64 increases when the agents act according to their subjective beliefs. Note also that the asset holdings at very high ages increase slightly, since old individuals are, on average, over-optimistic about their lifespan. However, this slight increase is still far from what we see in the data, a fact that we will return to in section 3.5.

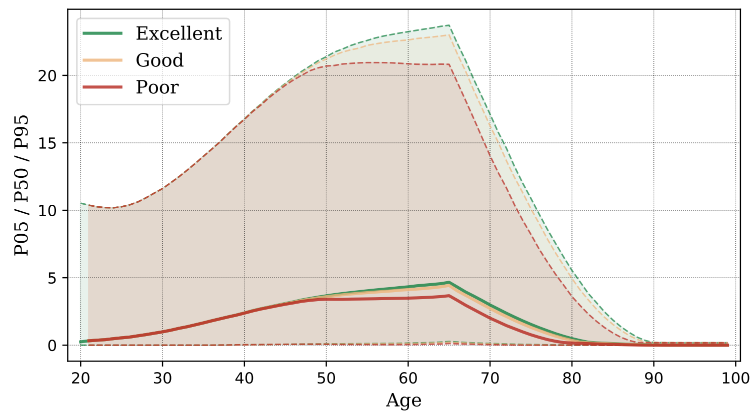
General equilibrium effects

In the previous section, we looked at the partial equilibrium responses from adding heterogeneity in survival expectations. We now move on to the effects on aggregate variables and the general equilibrium effects. The main takeaway is that overall, the general equilibrium effects of introducing survival heterogeneity are small. Table 3.4 shows three key statistics from the model.

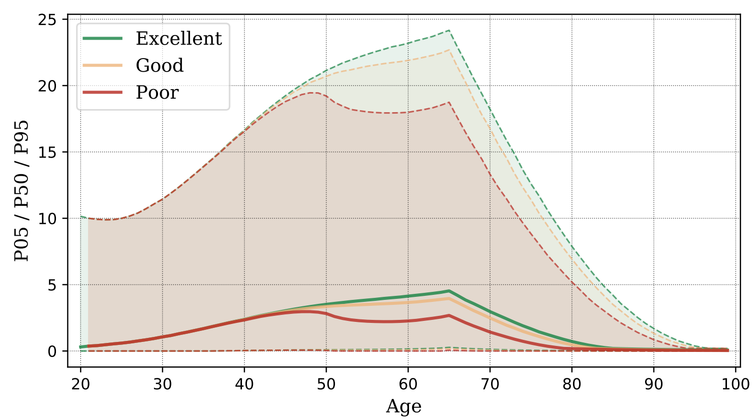
The reason why the interest rate increases slightly is the lower incentive to save on average among those in the ages 50 to 70. As Figure 3.11 shows, the demand for savings is substantially reduced at those ages among the individuals in bad health as compared to the baseline without heterogeneity. This effect is not counteracted by the slight increase in demand for savings by those in excellent health. Therefore, the interest rate needs to rise slightly to induce enough savings in the aggregate.



(a) No survival heterogeneity (NSH)



(b) Objective survival heterogeneity (OSH)



(c) Subjective survival heterogeneity (SSH)

Figure 3.13: Life-cycle profiles for wealth, model *without bequest motive*.

	Baseline: NSH (no heterogeneity)	OSH (objective exp.)	SSH (subjective exp.)
Interest rate	2.37%	2.38%	2.54%
Average bequest size	2.18	2.04	2.02
Wealth gini	0.65	0.65	0.65

Table 3.4: Comparing results from three scenarios (model with no bequest).

When agents act according to their subjective survival beliefs (SSH model), the interest rate increases slightly further, to 2.54%. The reason is that in the age group around 55–65, which is when people save the most, agents do, on average, have a downward bias in expected longevity. Hence, their incentives to save are reduced and consequently, the interest rate must increase to induce enough savings in the economy.

The next observation is that the average bequest is smaller in both the model with objective and the model with subjective survival beliefs, compared to the baseline model with no heterogeneity. Remember that in this model there is no bequest motive, so all bequests are accidental. Hence, with more information about the likelihood of their own demise, people can adjust their savings better and therefore the average accidental bequests are smaller in the model with objective survival heterogeneity.

When adding subjective survival heterogeneity, the accidental bequests are approximately of the same size as in the model with objective beliefs. The really old (aged 70+) are on average richer in the SSH model (due to their average over-optimism about survival), but the somewhat younger individuals (aged between 50 and 70) are somewhat poorer. These two effects cancel out and hence the average size of bequest is very similar in the two models with subjective and objective survival heterogeneity.

The last thing to note is that the wealth Gini is virtually unchanged between the three models. However, if we look at the within-cohort inequality, the picture is different. Figure 3.14 shows three statistics for wealth inequality, by age group: Gini, P90/P50 ratio, and P50/P10 ratio (the latter two in logs).

As can be seen, the difference is largest in the group of 90-94 year olds: within this age group, the Gini is 17pp higher in the model with subjective survival beliefs than in the baseline model. However, the two lower panels of the figure give more insight: in the baseline NHS model and in the OHS model, the P90/P50 metric is large, since the P50 is very small (approximately zero). However, in the SSH model, the P50 is small, but significantly larger than zero, so that the resulting P90/P50 ratio is still comparatively small. However, in the SSH model the P50/P10 ratio then becomes very large, since the P10 here is very small (approximately zero). In sum, the large swings in inequality measured for the very old age groups are driven by the fact that a smaller or larger share of the group has virtually zero assets, not that there are any

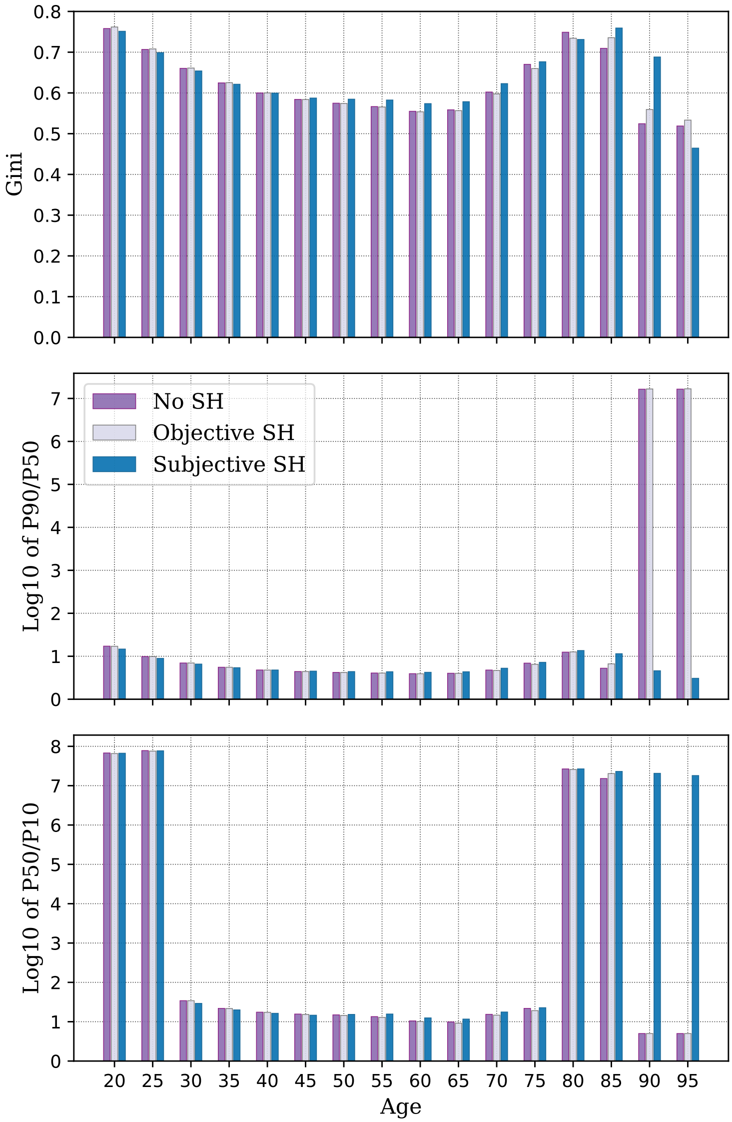


Figure 3.14: Wealth inequality in the model with no bequest motive, by 5-year age groups.

really rich individuals in this age group. The top 1% in the oldest age group in the OHS model owns 22% more assets than the top 1% in the same age group in the baseline NHS model. However, these asset levels are still low compared to the really asset rich in the middle aged groups.

To conclude, the within-cohort inequality within the very oldest cohorts is affected by the survival heterogeneity, as could be expected after seeing the large differences in the policy functions. However, the richest individuals in the economy are found in the age group 60–64, as we saw in the life-cycle profiles. For this age group, the "extra savings" by healthy individuals, due to a longer expected life, are very small, as we saw in Figure 3.11 – the main effect in this age group is that unhealthy individuals save less. Therefore, the effect of heterogeneity in survival on the top wealthiest individuals in the economy is very small, and therefore the effect on overall inequality is negligible.

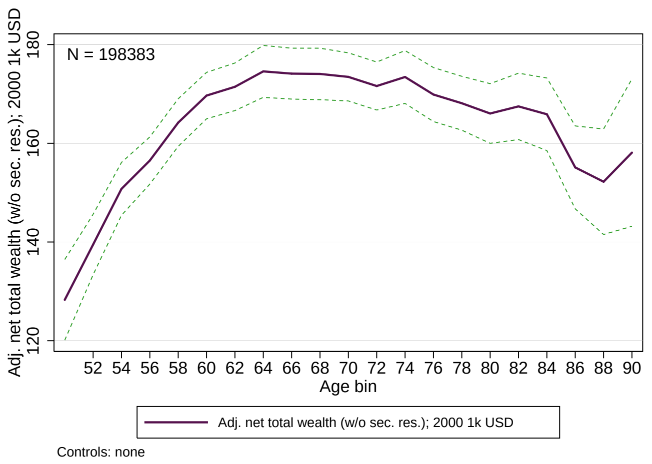
However, as previously noted, the asset levels among the oldest individuals in any of the scenarios analyzed in this section are far from the actual asset holding we see among older individuals. This motivates us to make the bequest motive operative, which is what we turn to in the next section.

Model with a bequest motive

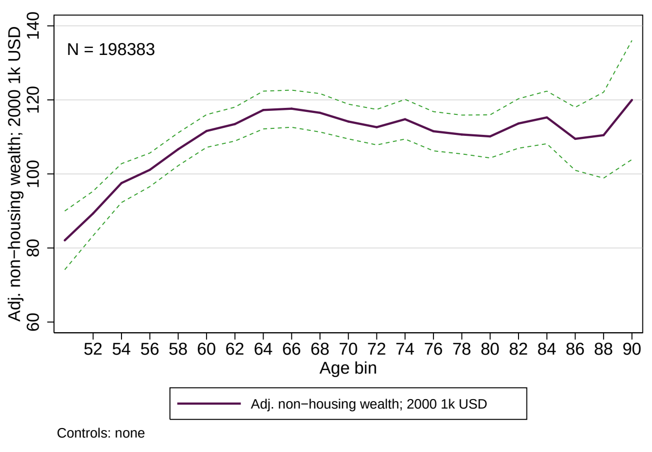
Figure 3.15 shows asset holdings in older ages, as measured by net total asset holdings including housing (net of mortgage), and net assets excluding housing and mortgages. As can be seen, the decumulation of assets in older ages is far from what a standard life-cycle model without a bequest motive predicts. Individuals keep a substantial amount of assets for reasons not present in the simplest life-cycle model. Motivated by this fact we add a bequest motive, as given in (3.8).

We calibrate the parameters of the bequest motive so that the median asset levels for the oldest age group, the bequests as a fraction of total wealth in the economy, and the fraction of taxed estates are approximately hit. The resulting life-cycle profile for the baseline NHS model is shown in Figure 3.16 and, as can be seen, the resulting asset holdings in older ages are now more in line with the data, with substantial asset holdings even for the average individual who survives beyond his/her 80th birthday. Note also that the peak wealth at the age of 65 is too pronounced as compared to the data, which is partly due to assuming a fixed retirement age.

We now move on to the effect of adding survival heterogeneity. Figure 3.17 shows the difference in the savings rate for the three scenarios. As can be seen, the results are counter-intuitive and perhaps also unexpected: individuals at the age of 50 in poor health save *more*, which is the complete opposite of the behavior we saw in the previous sections without bequest. The reason is as follows: since the agents in poor health are more likely to die soon, their effective discounting on the bequest utility decreases, and therefore the bequest motive for savings increases. Hence, the net effect is that these individuals want to save more.



(a) All assets.



(b) All assets, except net housing.

Figure 3.15: Mean asset holdings in older ages, by two-year age bins. 95% confidence intervals given by dashed green lines. Source: HRS.

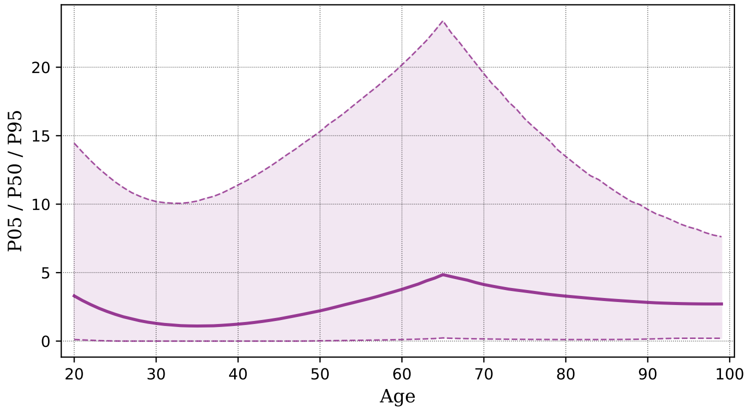


Figure 3.16: Life-cycle profile for wealth, baseline NHS model with a *standard bequest motive*.

The mechanism that a decrease in expected longevity increases savings due to the bequest motive is always present as long as we have a model with a bequest motive formulated as in (3.8). With a low weight on the bequest motive (lower than what we estimated the weight to be in line with the data), the life-expectancy channel could still dominate and the net effect could still be that individuals with a shorter life expectancy save less, but the underlying mechanism – that a decrease in expected longevity makes the agent want to save more for bequest reasons – would still be present.[14]

Figure 3.18 shows the resulting life-cycle profiles for the three scenarios. As can be seen, due to the bequest motive, individuals in older ages do not decumulate their wealth, and the resulting profile is more in line with the data. As is also clear, there is hardly any difference between individuals in poor vs. excellent health, neither in the model with objective survival heterogeneity, nor in the model with subjective heterogeneity. This is a direct result of what we saw in Figure 3.17: the savings rates are not that different as compared to the model with no survival heterogeneity.

As can be seen in the life-cycle profile for the SSH model, the median wealth among individuals in their 50s in poor health is slightly higher than among individuals in excellent health. This is the effect of the higher savings due to the bequest motive. However, wealth in the top P95 is slightly higher for 65-year-olds in excellent health than for 65-year-olds in poor health. The reason is that for extremely rich individuals, the savings motive to smooth consumption in the event of a longer life is becoming

14. In the appendix, section 3.A provides some more intuition and characterizes the required bequest weight in order to get a net effect of decreased savings in the event of an increased survival probability in a stylized two-period model.

relatively more important again, after having saved up so much that the marginal utility from the bequest is small.

3.6 Moving forward

The models used in the two preceding sections are in many ways the two most standard models one can think of: a canonical life-cycle model and a version extended with the standard bequest motive most commonly used in the macroeconomic literature.

The first model, the canonical life-cycle model without a bequest motive, gave rise to lower savings for individuals in bad health, and higher savings for individuals in good health. However, as is well known, the canonical life-cycle model gives rise to a counter-factually fast asset decumulation at older ages, contrary to the substantial asset holdings observed in the data.

This motivated us to include a bequest motive. Again, as is well known, such a model can be calibrated to fit the average asset profile in older ages reasonably well. However, this had counter-intuitive and perhaps unexpected implications for the savings behavior of individuals with different life expectancies. People in bad health, and hence a shorter expected life span, save *more* due to the lower effective discounting on the bequest motive. We argue that this mechanism is not plausible.

The conclusion is hence that neither of the standard models are really useful for understanding the true implications of survival heterogeneity on savings and wealth accumulation. This conclusion can be interpreted in two ways. Either there is no bequest motive, and the reason for savings in older ages is something else, or the current way of modelling the bequest motive is not the right one, at least not for analyzing the effect of differences in expected longevity. There is probably some truth in both statements.

There is overwhelming evidence that there is a bequest motive. This gives rise to the question of what a better way of formulating it would be. A natural first suggestion would be to try to shut down the "date-of-handover" channel for the bequest, by always discounting the utility of the bequest to some fixed future date, regardless of when the actual death occurs. However, in practice, this is isomorphic to having a time shifting weight on the bequest, and does not solve the problem.[15]

Another suggestion would be to not discount the bequest at all. One could argue that for the realized utility of the bequest, it does not matter at all exactly when the

15. For instance, one could model the bequest motive as $\beta^{T-t}\theta_B\frac{(\kappa+a)^{1-\sigma}}{1-\sigma}$, which means that the agent always discounts the bequest utility as if it were handed over in time T. However, the formulation is isomorphic to a bequest motive of the standard type, but with a weight that is increasing over time. This gives rise to counter-factually large savings late in life, and does not remove the expected-date-of-handover effect.

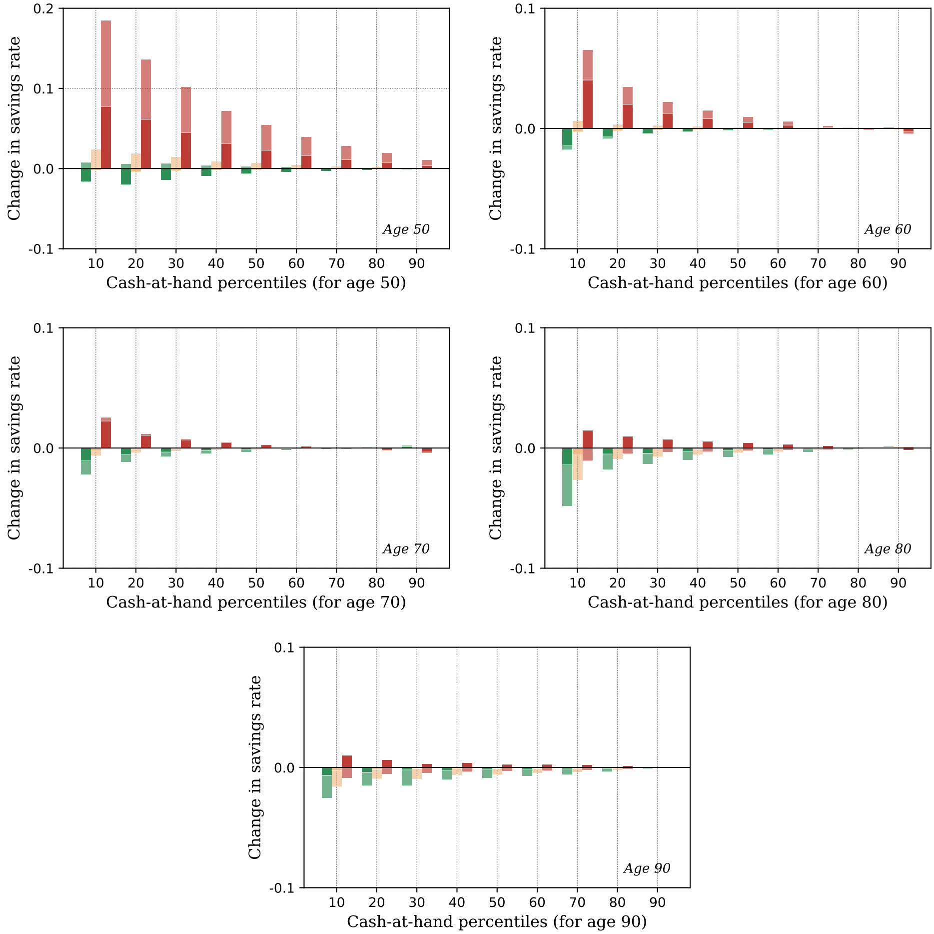
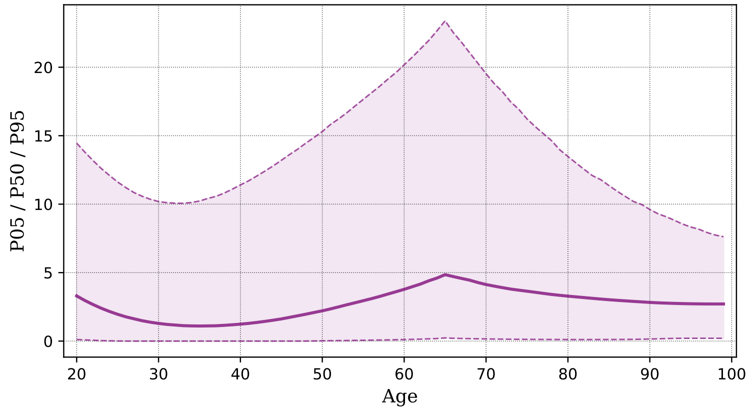
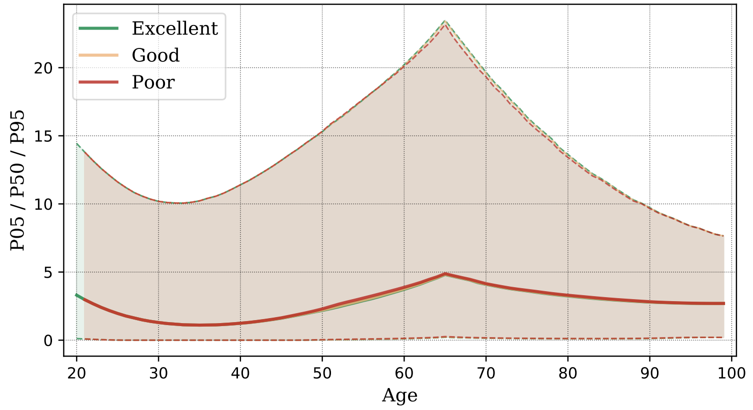


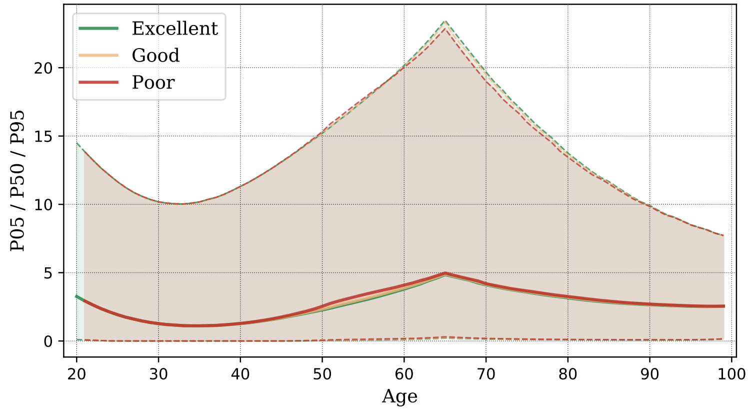
Figure 3.17: Difference in the savings rate for the model with a *standard bequest motive* (percentage points/100, i.e.,, 0.1 indicates an difference of 10 pp). The darker color shows the model with objective survival heterogeneity compared to the baseline model with no survival heterogeneity. The lighter color shows the additional effect of adding subjective survival beliefs. The x-axis depicts cash-at-hand percentiles by age (equilibrium values for the baseline model). The color indicates health state: green is excellent while red is poor health.



(a) No survival heterogeneity (NSH)



(b) Objective survival heterogeneity (OSH)



(c) Subjective survival heterogeneity (SSH)

Figure 3.18: Resulting life-cycle profiles for wealth, model with a *standard bequest motive*.

wealth is handed over, rather it is only the sum that matters. However, a simple formulation with no discounting of the bequest term in the utility function gives rise to dynamically inconsistent preferences.

One might think that the inclusion of other economic effects related to the current health status would remedy the problem. For instance, if unhealthy individuals were subject to lower wages (as they are in the data), expensive medical shocks and/or had some inherent low-savings characteristics, they would have lower asset holdings on average. Hence, since leaving a bequest is a luxury good, they would not, on average, have a very strong bequest motive, and therefore they would on average save less. This argument is true in the sense that it would be a remedy for some of the counterfactual cross-sectional implications: we would most certainly get a model where the unhealthy individuals again have lower asset holdings, as in the data. However, the implausible behavioral mechanism would still exist. Imagine for instance a rich person who, thanks to good income and no large negative shocks, saved up a lot. For this asset rich person, the bequest motive is operational. Now imagine that this person receives a bad health shock (think about a cancer diagnosis), which shortens his/her expected life span. Then, the implication for this person's savings behavior would be that he/she saves more, despite the shorter expected life span, due to the bequest motive. Hence, we argue that even though more health-related shocks could be a remedy for the cross-sectional counterfactual health-wealth gradient, it would not solve the underlying implausible connection between survival expectancy and savings rate.

A more promising path is to explicitly incorporate the parent-child connection, allow for inter-vivo bequests, and maximize over the dynasty utility. In such a model, an increase in the expected life span would translate into an expectation of a longer period with two agents alive in parallel in the dynasty. However, then it would be important to also have a motive for asset holdings in older ages for childless individuals.[16] With a warm-glow motive, the exact definition of who eventually gets the bequest is not as important.

16. According to Livingston 2015, the number of childless women between 40 and 44 has varied between 15 and 20 percent during the last 30 years.

Bibliography

Brinca, Pedro, Hans A. Holter, Per Krusell, and Laurence Malafry. 2016. Fiscal multipliers in the 21st century. “Inequality, Institutions, and Redistribution”held at the Stern School of Business,New York University,April 24-25, 2015, *Journal of Monetary Economics* 77:53–69.

Cocco, João F., Francisco J. Gomes, and Pascal J. Maenhout. 2005. Consumption and portfolio choice over the life cycle. *The Review of Financial Studies* 18 (2): 491–533.

Coile, Courtney, and Kevin Milligan. 2009. How household portfolios evolve after retirement: the effect of aging and health shocks. *Review of Income and Wealth* 55 (2): 226–248.

De Nardi, Mariacristina. 2004. Wealth inequality and intergenerational links. *The Review of Economic Studies* 71 (3): 743–768.

De Nardi, Mariacristina, and Giulio Fella. 2017. Saving and wealth inequality. *Review of Economic Dynamics* 26:280–300.

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. *The lifetime costs of bad health.* Technical report. National Bureau of Economic Research.

Foltyn, Richard, and Jonna Olsson. 2019. *Health dynamics and heterogeneous life expectancies.* Technical report. Working paper, Institute for International Economic Studies.

Gale, William G, and John Karl Scholz. 1994. Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives* 8 (4): 145–160.

Groneck, Max, Alexander Ludwig, and Alexander Zimper. 2016. A life-cycle model with ambiguous survival beliefs. *Journal of Economic Theory* 162:137–180.

———. 2017. The impact of biases in survival beliefs on savings behavior.

Hamermesh, Daniel S. 1985. Expectations, life expectancy, and economic behavior. *The Quarterly Journal of Economics* 100 (2): 389–408.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. Optimal tax progressivity: an analytical framework. *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Heimer, Rawley Z., Kristian Ove R. Myrseth, and Raphael S. Schoenle. 2019. Yolo: mortality beliefs and household finance puzzles. *Journal of Finance* 74 (6): 2957–2996.

Hendricks, Lutz. 2007. How important is discount rate heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control* 31 (9): 3042–3068.

Hurd, Michael D., and Kathleen McGarry. 2002. The predictive validity of subjective probabilities of survival. *Economic Journal* 112 (482): 966–985.

Hurd, Michael D, James P Smith, and Julie M Zissimopoulos. 2004. The effects of subjective survival on retirement and social security claiming. *Journal of Applied Econometrics* 19 (6): 761–775.

Krueger, D., K. Mitman, and F. Perri. 2016. Chapter 11 - macroeconomics and household heterogeneity, edited by John B. Taylor and Harald Uhlig, 2:843–921. Handbook of Macroeconomics. Elsevier.

Krusell, Per, and Anthony A. Smith Jr. 1998. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy* 106 (5): 867–896.

Lee, Jinkook, and Hyungsoo Kim. 2008. A longitudinal analysis of the impact of health shocks on the wealth of elders. *Journal of Population Economics* 21 (1): 217–230.

Livingston, Gretchen. 2015. *Childlessness falls, family size grows among highly educated mothers.* Pew Research Center.

Ludwig, Alexander, and Alexander Zimper. 2013. A parsimonious model of subjective life expectancy. *Theory and Decision* 75 (4): 519–541.

Pijoan-Mas, Josep, and José-Víctor Ríos-Rull. 2014. Heterogeneity in expected longevities. *Demography* 51 (6): 2075–2102.

Poterba, James M., Steven F. Venti, and David A. Wise. 2017. The asset cost of poor health. *The Journal of the Economics of Ageing* 9:172–184.

Smith, James P. 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives* 13 (2): 145–166.

Smith, V. Kerry, Donald H. Taylor, and Frank A. Sloan. 2001. Longevity expectations and death: can people predict their own demise? *The American Economic Review* 91 (4): 1126–1134.

Appendix

3.A More on the bequest motive

With the conventional way of formulating a bequest motive, the effect of an increase in survival has unexpected and non-intuitive effects on savings. To illustrate the argument, we write down a simple two-period model where the only uncertainty is the survival probability, π, between the first and the second period. After the second period, the agent dies with certainty. We assume that the agent has some initial assets a_0, the discount factor is $\beta < 1$, and the gross interest rate is R. Hence, the agent solves

$$\max_{c_0, c_1} \left\{ u(c_0) + \beta \Big(\pi \Big[u(c_1) + \beta V_b(a_2) \Big] + (1 - \pi) V_b(a_1) \Big) \right\} \tag{3.9}$$

subject to the constraints

$$\begin{aligned} c_t &> 0 \quad \forall t \\ a_1 &= a_0 - c_0, \quad a_2 = R a_1 - c_1. \end{aligned}$$

Taking first-order conditions with respect to the choice variables gives the following optimality conditions:

$$\begin{aligned} \frac{\partial u}{\partial c_0} &= \beta^2 \pi R \left. \frac{\partial V_b}{\partial a_2} \right|_{a_2^*} + \beta(1 - \pi) \left. \frac{\partial V_b}{\partial a_1} \right|_{a_1^*} \\ \frac{\partial u}{\partial c_1} &= \beta \left. \frac{\partial V_b}{\partial a_2} \right|_{a_2^*} \end{aligned}$$

where a_1^* and a_2^* denote the optimal choice of savings in each period. Hence, the impact of an increase in survival probability on c_0^*, optimal first period consumption, is ambiguous and depends on how the bequest motive is parametrized.

The utility function u has the usual CRRA functional form:

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma}$$

and we use the standard bequest motive:

$$V_b(a) = \theta_B \frac{(a + \kappa)^{1-\sigma} - 1}{1 - \sigma} \tag{3.10}$$

where θ_B determines the strength of the bequest motive, and κ determines to what extent bequests are a luxury good. For simplicity, we assume $\kappa = 0$ and solve for the agent's problem as given in (3.9). Some algebra gives that if an increase in the survival probability makes the agent consume more or less in the first period, i.e., the sign of $\frac{\partial c_0^*}{\partial \pi}$, is determined by the ratio

$$\theta_B \lesseqgtr \left(\frac{R^{\frac{1-\sigma}{\sigma}}}{1 - R^{\frac{1-\sigma}{\sigma}} \beta^{1/\sigma}} \right)^{\sigma} \equiv \widehat{\theta}_B \tag{3.11}$$

If $\theta_B < \widehat{\theta}_B$, an increase in the survival probability has the expected effect: the agent consumes less in the current period and saves more, given that it is more likely to survive to the next period.

On the other hand, if $\theta_B > \widehat{\theta}_B$, an increase in the survival probability leads to *decreased* savings and more consumption in the first period. There are two mechanisms behind this. First, when the probability of surviving increases, the effective discounting of the next-period bequest utility increases, and hence the incentive to save decreases. We call this the expected-date-of-handover channel. Second, an increase in the survival probability leads to a higher expected interest rate income over the remaining life, and therefore the agent can afford more consumption also in the first period. We call this channel the income effect. If the weight on the bequest motive is high, these two effects dominate the effect of wanting to save for a longer expected life.

In the above example, we assumed $\kappa = 0$. It can be shown that if we assume $R = 1$, the sign of $\frac{\partial c_0^*}{\partial \pi}$ is actually independent of κ (although it still affects the level of savings). However, if we allow for a positive interest rate, the extent to which the bequest is a luxury good in combination with the level of initial assets, a_0, matters. Moreover, we assumed no second-period income in our above example. If we allow for income in the second period, the level of that income compared to the initial assets affects the incentives. For a given a_0, the higher the second-period income, the less bequest weight is needed to get $\frac{\partial c_0^*}{\partial \pi} > 0$ – since then the bequest becomes relatively more important as a reason to save.

3.B Social Security system

Retirement benefits

First, consider the following stylized version of the actual retirement income formula used in the U.S. social security system, where $\overline{e}$ is an (annualized) measure of historical average monthly earnings, $b_1^{\$}$ and $b_2^{\$}$ are bend points in USD for some reference year, and $e_{max}^{\$}$ is the contribution and benefit base (CBB), i.e., the maximum

earnings subject to payroll taxes. Retirement income measured in USD, $\iota_R^\$$, is then approximately given by

$$\iota_R^\$(\overline{e}) = \begin{cases} \rho_1\overline{e} & \text{if } \overline{e} \leq b_1^\$ \\ \rho_1 b_1^\$ + \rho_2\left(\overline{e} - b_1^\$\right) & \text{if } b_1^\$ < \overline{e} \leq b_2^\$ \\ \rho_1 b_1^\$ + \rho_2\left(b_2^\$ - b_1^\$\right) + \rho_3\left(\min\left\{e_{\max}^\$, \overline{e}\right\} - b_2^\$\right) & \text{else} \end{cases}$$

where ρ_1, ρ_2 and ρ_3 are decreasing replacement rates applied to earnings ranges bracketed by the bend points $b_i^\$$.

In the model, we define retirement benefits to be a product of the following components:

$$\iota_R(p) = w \times y_R(p) = w \times \omega_{T_R-1}\overline{\epsilon} p_R(p)$$

where the function $p_R(\cdot)$ is constructed below analogously to $\iota_R^\$$.

To this end, denote by $e_{med}^\$$ the USD median earnings in the reference year. To express the bend points in terms of the persistent labor state in retirement, we implicitly define $\widetilde{p}_i$ corresponding to $b_i^\$$ as

$$\frac{b_i^\$}{e_{med}^\$} = \frac{w \times \omega_{T_R-1}\widetilde{p}_i\overline{\epsilon}}{w \times y_{med}}$$

where we have normalized the bend point $b_i^\$$ by real-world median earnings and the model counterpart with the median earnings in the model. We find that

$$\widetilde{p}_i = \frac{\left(b_i^\$/e_{med}^\$\right) y_{med}}{\omega_{T_R-1}\overline{\epsilon}}.$$

Analogously, the CBB in terms of persistent labor productivity is

$$\widetilde{p}_{max} = \frac{\left(e_{max}^\$/e_{med}^\$\right) y_{med}}{\omega_{T_R-1}\overline{\epsilon}}.$$

By factoring out the common term $\omega_{T_R-1}\overline{\epsilon}w$ that is independent of a retired individual's idiosyncratic state vector, we can write the replacement formula purely in terms of the permanent labor state as follows:

$$p_R(p) = \begin{cases} \rho_1 p & \text{if } p \leq \widetilde{p}_1 \\ \rho_1\widetilde{p}_1 + \rho_2\left(p - \widetilde{p}_1\right) & \text{if } \widetilde{p}_1 < p \leq \widetilde{p}_2 \\ \rho_1\widetilde{p}_1 + \rho_2\left(\widetilde{p}_2 - \widetilde{p}_1\right) + \rho_3\left(\min\left\{\widetilde{p}_{max}, p\right\} - \widetilde{p}_2\right) & \text{else} \end{cases}$$

Note that this implies that the bend points and the CBB are proportional to the wage level w.

The government expenditures on retirement benefits in a period are given by

$$\begin{aligned} G_{ss} &= \sum_{t=T_R}^{N_t} \sum_p \mu_t(t)\mu_p(p)\iota_R(p) \\ &= w\omega_{T_R-1}\bar{\epsilon} \cdot \overline{p}_R \Pi_R \end{aligned} \tag{3.12}$$

which is a weighted sum over the the retirement incomes received by all retired cohorts, with weights μ_t and μ_p denoting the PMFs of the ergodic distribution of age and persistent labor states, respectively. We denote the mass of retired individuals by

$$\Pi_R = \sum_{t=T_R}^{N_t} \mu_t(t)$$

and the average permanent component of retirement income as

$$\overline{p}_R = \sum_p \mu_p(p) p_R(p).$$

Social security budget balance

The payroll taxes raised each period are

$$\overline{T}_{ss} = \sum_{t=1}^{T_R-1} \sum_p \sum_\epsilon \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon) T_{ss}\left(y_e\right) w \tag{3.13}$$

where $\mu_\epsilon(\epsilon)$ is the ergodic distribution over transitory labor shocks. The payroll tax function is defined as

$$T_{ss}\left(y\right) = \tau_{ss} \times \min\{y_{max}, y\}$$

where y_{max} expresses maximum taxable earnings in terms of labor productivity, i.e.,

$$y_{max} = \left(e^{\$}_{max}/e^{\$}_{med}\right) y_{med}.$$

To balance the social security system, we need to find τ_{ss} such that $G_{ss} = \overline{T}_{ss}$. Equating $G_{ss} = \overline{T}_{ss}$ implies that

$$\tau_{ss} = \frac{\omega_{T_R-1}\overline{p}_R\bar{\epsilon}\Pi_R}{\sum_{t=1}^{T_R-1} \sum_p \sum_\epsilon \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon) \min\left\{y_{max}, \omega_t p \epsilon\right\}}. \tag{3.14}$$

3.C Government budget

Estate taxes

Our solution algorithm requires a continuously differentiable tax rate on estates. Additionally, we want to impose a tax exemption to make sure that only a small fraction of estates is subject to the estate tax.

To this end, we use the cosine function to create a marginal tax rate that is defined as follows:

$$\frac{\partial T_b(b)}{\partial b} = \begin{cases} 0 & \text{if } b \leq \chi_b \\ \frac{\tau_b}{2}\left[\cos\left(\pi\left[\frac{b-\chi_b}{B} - 1\right]\right) + 1\right] & \text{if } \chi_b < b \leq \chi_b + B \\ \tau_b & \text{else} \end{cases} \tag{3.15}$$

We assume that the marginal tax rate is increasing on the interval $[\chi_b, \chi_b + B]$ and constant everywhere else. The tax function itself is obtained by integrating (3.15), which gives

$$T_b(b) = \begin{cases} 0 & \text{if } b \leq \chi_b \\ \frac{\tau_b}{2}\left[\sin\left(\pi\left[\frac{b-\chi_b}{B} - 1\right]\right)\frac{B}{\pi} + b - \chi_b\right] & \text{if } \chi_b < b \leq \chi_b + B \\ \tau_b(b - \chi_b - B) + \frac{\tau_b}{2}B & \text{else} \end{cases}$$

Income taxes

In this section we derive an expression for the total amount of income taxes raised by the government. Before proceeding, we state the following useful definitions: We denote by $\overline{p}$ the average persistent labor shock,

$$\overline{p} = \sum_p \mu_p(p)p \tag{3.16}$$

and by Π_{LF} the size of the labor force,

$$\Pi_{LF} = \sum_{t=1}^{T_R-1} \mu_t(t) = 1 - \Pi_R$$

Additionally, average labor productivity can be defined as

$$\overline{y} = \Pi_{LF}^{-1}\left[\sum_t^{T_R-1}\sum_p\sum_\epsilon \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon)\omega_t p\epsilon\right] = \Pi_{LF}^{-1}\left[\overline{p}\cdot\overline{\epsilon}\sum_h^{T_R-1}\mu_t(t)\omega_t\right].$$

Now, consider the aggregate tax revenues raised from working individuals, which are given by

$$
\begin{aligned}
T_e &= \sum_{t=1}^{T_R-1} \sum_p \sum_\epsilon \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon)\Big[\big(\omega_t p\epsilon - T_{ss}\left(\omega_h p\epsilon\right)\big)w \\
&\qquad - \lambda\Big(\big(\omega_t p\epsilon - T_{ss}\left(\omega_t p\epsilon\right)\big)w\Big)^{1-\tau}\Big] \\
&= \Big[w\Pi_{LF}\overline{y} - \lambda w^{1-\tau}\overline{y}_{-ss,\tau}\Big] - \overline{T}_{ss}
\end{aligned}
$$

where we define

$$
\overline{y}_{-ss,\tau} = \sum_{t=1}^{T_R-1} \sum_p \sum_\epsilon \mu_t(t)\mu_p(p)\mu_\epsilon(\epsilon)\Big(\omega_t p\epsilon - T_{ss}\left(\omega_t p\epsilon\right)\Big)^{1-\tau}
$$

to simplify the notation.

Income taxes raised from retired individuals amount to

$$
\begin{aligned}
T_R &= \sum_{t=T_R}^{N_t} \sum_p \mu_t(t)\mu_p(p)\left[y_R(p)w - \lambda\Big(y_R(p)w\Big)^{1-\tau}\right] \\
&= \sum_{t=T_R}^{N_t} \sum_p \mu_t(t)\mu_p(p)\left[\omega_{T_R-1}p_R(p)\overline{\epsilon}w - \lambda\Big(\omega_{T_R-1}p_R(p)\overline{\epsilon}w\Big)^{1-\tau}\right] \\
&= \Pi_R\left[w\omega_{T_R-1}\overline{p}_R\overline{\epsilon} - \lambda w^{1-\tau}\omega_{T_R-1}^{1-\tau}\overline{p}_{R,\tau}\overline{\epsilon}^{1-\tau}\right] \\
&= \overline{T}_{ss} - \lambda\Pi_R w^{1-\tau}\omega_{T_R-1}^{1-\tau}\overline{p}_{R,\tau}\overline{\epsilon}^{1-\tau}
\end{aligned}
$$

with

$$
\overline{p}_{R,\tau} = \sum_p \mu_p(p)p_R(p)^{1-\tau}
$$

Thus, the total revenue from income taxes is

$$
\begin{aligned}
T_{inc} &= T_e + T_R \\
&= w\Pi_{LF}\overline{y} - \lambda w^{1-\tau}\Big[\overline{y}_{-ss,\tau} + \Pi_R\omega_{T_R-1}^{1-\tau}\overline{p}_{R,\tau}\overline{\epsilon}^{1-\tau}\Big]
\end{aligned}
$$

Chapter 4

On the Redistributive Effects of Government Bailouts in the Mortgage Market

(joint with Dirk Krueger and Kurt Mitman)

4.1 Introduction

The spectacular boom and bust in house prices in the 2000s in the U.S. was a key driver of the Great Recession: the bust in house prices triggered a collapse in aggregate demand and a banking crisis (Mian and Sufi 2014; Mian, Rao, and Sufi 2013). The government's response to the financial crisis implied substantial bailouts of the mortgage market. While there was no formal bailout guarantee to the so-called Government Sponsored Enterprises (GSEs, e.g., Fannie Mae and Freddie Mae) before the crisis, investors treated GSE debt as if there was a bailout guarantee in place. The CBO (2001) noted that "the implicit federal guarantee leads investors in a GSE debt or MBS to believe that the federal government bears most if not all the risk." This implicit guarantee became explicit in September 2008 when the U.S. government took the GSEs into conservatorship and committed $180 billion to help the GSEs remain solvent (Frame 2010).

In this paper, we explore the distributional and macroeconomic consequences of the government guarantees for the GSEs. We examine the implications of the bailout through the lens of a general-equilibrium, incomplete-markets model of the U.S. economy that includes detailed modeling of the housing finance sector and generates realistic distributions of consumption, savings, home-ownership and default among households. Households are subject to idiosyncratic shocks to their labor earnings, and home-owners face idiosyncratic shocks to the value of their homes. In addition to idiosyncratic risk, the economy is subject to aggregate risk. We think of the economy as being in one of three idiosyncratic states: boom, mild recession, and severe recession. The aggregate states correspond to movements in total-factor productivity (TFP) and the (exogenous) aggregate supply of labor. Endogenous propagation in

the economy occurs via household adjustments in aggregate investment in physical and housing capital.

Financial intermediates face a bailout guarantee from the government in severe economic downturns. As a result, the expected loss given default in bad times is not passed through to consumers in the form of higher interest rates. Under the bailout policy, households can, therefore, take on larger mortgages at lower interest rates than absent the bailout. We find that compared to a world without bailouts, more low-income households become home-owners. As a result, gross housing investment, the housing stock, and house prices are higher in the economy with bailouts, leading to a potential misallocation of resources to housing away from capital used in production. The ultimate goal of the paper is first to quantify the positive effects of the bailout policy, in terms of its affect on aggregates and cross-sectional distributions across households. Finally, we plan to conduct a welfare analysis of the bailout policy, to understand which households gain or lose from the bailout policy.

Related Literature

Our paper relates to several strands of the quantitative macroeconomics literature. Our incomplete-markets general equilibrium framework builds on the pioneering work of Aiyagari (1994) and Krusell and Smith (1998). We introduce housing and a household finance sector along the lines of Favilukis, Ludvigson, and Van Nieuwerburgh (2017). However, in contrast to their innovative study, we explicitly allow for household default on mortgages and thus provide a role for government bailouts in the mortgage sector. A separate and related literature has focused on understanding the causes of the boom and bust in house prices during this episode (e.g., Kaplan, Mitman, and Violante (2017) and Justiniano, Primiceri, and Tambalotti (2017)).[1]

Our paper also contributes toward the literature that studies the positive and normative effects of government interventions in the housing market. Beginning with the work of Gervais (2002), a long literature has investigated on the effects of the mortgage interest deduction and its potential repeal on home-ownership, house prices and welfare (Karlman, Kinnerud, and Kragh-Sorensen 2020; Sommer and Sullivan 2018). Closest in question to our paper is the pioneering work by Elenev, Landvoigt, and Van Nieuwerburgh (2016) who also study the consequences of removing bailout guarantees. In contrast to our study, they focus more on the financial sector and abstract from idiosyncratic risk. We view the two papers as complementary, with theirs a more detailed focus on the consequences of bailouts for risk-averse bankers,

1. While our study is related to that literature in that we need to generate a boom and bust in house prices to study our policy question we don't see our ability to generate such a boom and bust as a standalone contribution of the paper. For a broader overview of the housing and macro literature, we refer the interested reader to the excellent review by Piazzesi, Schneider, and Tuzel (2007).

and ours more focused on the distributional consequences of bailouts for households and aggregates.

4.2 The model

Stochastic structure

Our model features aggregate productivity risk as well as idiosyncratic income and house price risk. It can be seen as an extension of the environment in Krusell and Smith (1997, 1998) to a model with risky real estate, short-term mortgages and a foreclosure option, as in the stationary model of Jeske, Krueger, and Mitman (2013). In order to induce business cycles, the economy is subject to aggregate productivity shocks, denoted by $z \in \{z_\ell, z_m, z_h\}$. We assume $z_\ell \ll z_m < z_h$, with the hope of generating something like a slowish boom and severe bust in house prices. The process follows a finite Markov chain with transition matrix $\pi(z'|z)$.

We interpret z_h as normal times, z_m as normal recessions and z_ℓ as a severe economic crisis, such as the great depression or the great recession. The aggregate shock z affects the aggregate component of TFP in the final goods sector $A(z)$, the aggregate labor endowment $L(z)$, and the distribution of idiosyncratic house price risk shocks $\delta \sim F_z(\delta)$.

In addition, households face uninsurable idiosyncratic income risk that we will spell out in greater detail below. As a consequence the endogenous aggregate state variables, from now on summarized by S, include the cross-sectional joint income and wealth (or more precisely, cash at hand) distribution across households. We denote the aggregate law of motion as perceived by agents in this economy by $S' = \Gamma(z, S, z')$. In what follows we directly set up the economy recursively, thereby sidestepping the sequential formulation of the model.

Financial markets

In our model, households can trade a full set of Arrow securities that pay out contingent on the *aggregate* state of the world. We denote the price of an Arrow security which pays one unit of of consumption tomorrow conditional on a realization z' by $P_b(z, S; z')$. The price of these bonds will be determined in equilibrium. Note that the markets for individual income, survival and house price risk are assumed to be incomplete, and hence households are unable to perfectly insure against idiosyncratic risk.

The motivation for allowing trade in Arrow securities contingent on the aggregate state is that it allows the firms in our model to diversify away risk as they do not face any idiosyncratic uncertainty. This enables us to set up representative firms that make zero profits in all states of the world, and hence we can ignore the issue of firm

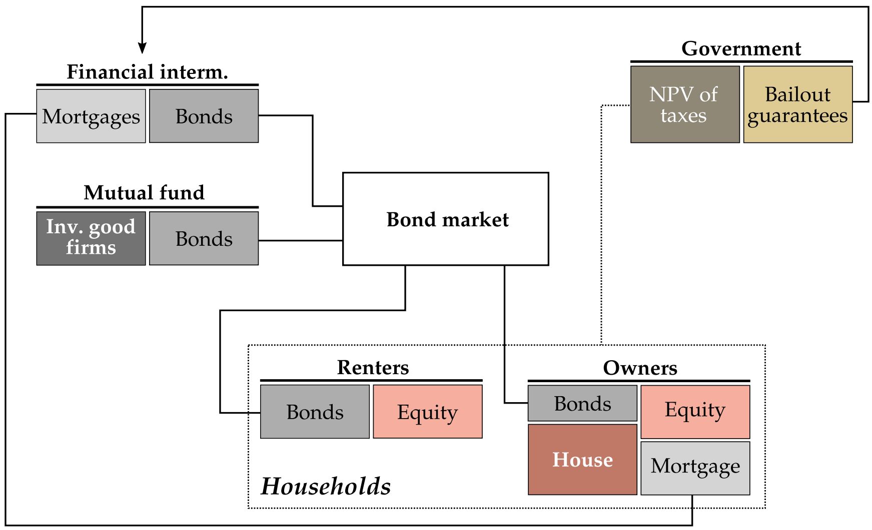


Figure 4.1: High-level overview of selected agents and their balance sheets

ownership.[2] We need households to participate in Arrow bond markets in order to have a counterparty for the bonds issued by various intermediaries, which we discuss in more detail below.

In Figure 4.1 we provide a high-level overview of the model for a subset of agents and their balance sheet. We omit the housing market to keep the illustration simple.

Final output production technology

Final output is produced by a representative, perfectly competitive firm that operates a constant returns to scale technology

$$Y = A(z)K^{\alpha}L(z)^{1-\alpha}$$

where Y is the final goods output and K and L are the capital and labor inputs the firm rents from perfectly competitive factor markets at rental rates $w(z,S)$ and $r(z,S)$, respectively. For future reference, these will be given in equilibrium as

$$w(z,S) = (1-\alpha)\frac{Y(z,S)}{L(z)}$$

$$r(z,S) = \alpha\frac{Y(z,S)}{K(z)}$$

2. As we discuss below, this is not true for the real estate construction firm in the current iteration of the model, but will be addressed in a future version.

where we have already anticipated the labor market clearing condition equating labor demand to exogenously given aggregate labor supply $L(z)$.

Investment goods producers

A representative investment goods producer owns a capital stock k, rents it out to the final output-producing firms at a rental rate $r(z,S)$, makes investment decisions and faces a capital adjustment cost.[3] In order to install new investment i the firm has to purchase final goods in the amount of $i + \psi_k\left(\frac{i}{k}\right)k$, where the function $\psi_k(\bullet)$ satisfies the following

Assumption 1. The function $\psi_k(\bullet)$ is twice differentiable, strictly increasing and strictly convex. The function $\psi_k\left(\frac{i}{k}\right)k$ is homogeneous of degree 1 in i, k.

Example 1. An example that satisfies this assumption is $\psi_k\left(\frac{i}{k}\right) = \frac{1}{2}\overline{\psi}_k\left(\frac{i}{k} - \delta_k\right)^2$ where $\overline{\psi}_k \geq 0$ is a parameter.

The recursive problem of the firm is then given by

$$\begin{aligned}
V_k(k;z,S) &= \max_{i,k'}\left\{d(k;z,S) + \sum_{z'} P_b(z,S;z')\, V_k(k';z',S')\right\} \\
\text{s.t.} \quad & k' = (1-\delta_k)k + i \qquad (4.1) \\
& k' \geq 0 \\
& S' = \Gamma(z,S,z')
\end{aligned}$$

where V_k represents the cum-dividend value of the firm, (4.1) is the standard capital accumulation equation, and current dividends of the firm are given by

$$\begin{aligned}
d(k;z,S) &= r(z,S)\,k - i - \psi_k\left(\frac{i}{k}\right)k \\
&= r(z,S)\,k - k' + (1-\delta_k)k - \psi_k\left(\frac{i}{k}\right)k \\
&= \left[r(z,S) + (1-\delta_k) - \frac{k'}{k} - \psi_k\left(\frac{i}{k}\right)\right]k
\end{aligned}$$

The firm values cash flows in next period's state z' using the price of an Arrow security which pays one unit of consumption tomorrow contingent on z'.

Similar to Hayashi (1982), Assumption 1 and the fact that the firm takes the rental rate of capital as given gives rise to

3. We denote firm-specific variables in lower case. Thus a firm owns a capital stock k, but the rental rate depends on the aggregate capital stock K in the economy, which is included in the aggregate state S. Since the firm is representative we will later set $k = K$.

Proposition 1. The value function $V_k(k; z, S)$ and the optimal investment policy function are linear in k:

$$V_k(k; z, S) = \kappa_V(z, S) \cdot k$$
$$i(k; z, S) = \kappa_i(z, S) \cdot k$$

For an outline of a proof we refer to section 4.A in the appendix. Since the firm problem is linear in k, without loss of generality we assume that there is a single representative investment goods producer that holds the aggregate capital stock K.

Mutual fund

Each period, a representative mutual fund purchases all stocks of the representative investment goods producer and issues Arrow securities $B_k(z, S; z')$ for each z' to do so. At the beginning of next period, it receives all dividends $d(K'; z, S)$ and sells off all shares to liquidate its position of the Arrow securities. The fund's budget constraint today reads

$$\sum_{z'} P_b(z, S; z') B_k(z, S; z') = V_k^e(K'; z, S)$$

where the right-hand side is the ex-dividend value of the representative investment goods producer. For the mutual fund to make zero profits, the amount of Arrow bonds it has to issue contingent on each z' must be identical to the cum-dividend value of the investment goods producing firm in that state, i.e.,

$$B_k(z, S; z') = V_k(K'; z, S) \qquad \forall\, z'$$

Real estate production technology

Perfectly divisible houses are produced by a real estate construction company.[4] The firm has access to a technology that uses final goods C_h for the production of residential fixed investment. The production function for residential fixed investment reads as:

$$I_h = A_h(z) C_h^{1-\alpha_h}$$

where $\alpha_h \in (0, 1)$ is a technology parameter. In addition, to produce I_h new houses the firm has to pay a convex adjustment cost $\phi_h\left(\frac{I_h}{H}\right) H$, where H is the aggregate housing stock in the economy and I_h is aggregate housing construction. The real

4. For now we assume that the real estate construction firm takes the price of housing P_h as given. In the next version of this paper we will adopt the setup in Favilukis, Ludvigson, and Van Nieuwerburgh (2017) where government-issued permits enter the firm's production function such that the problem admits a representative firm.

estate construction company takes prices of new houses $P_h(z,S)$ as well as the aggregate housing as given and maximizes

$$\begin{aligned}\Pi_h(z,S) &= \max_{I_h,C_h\geq 0} P_h(z,S)\, I_h - \psi_h\left(\frac{I_h}{H}\right) H - C_h \qquad (4.2)\\ \text{s.t.} \qquad I_h &= A_h(z) C_h^{1-\alpha_h}\end{aligned}$$

Under these assumptions the real estate construction company makes non-zero profits which are distributed uniformly to all households in a lump-sum fashion.

Real estate rental company

There is a representative rental company that, in the current period, buys real estate H_r for unit price $P_h(z,S)$, rents it out immediately at rental rate $P_\ell(z,S)$ and tomorrow sells off the undepreciated part of the housing stock for price $P_h(z',S')$. Profits of the real estate rental company are given by

$$\Pi_r = \max_{H_r\geq 0} H_r\left[P_\ell(z,S) + \sum_{z'} P_b\left(z,S;z'\right)\left(1-\overline{\delta}\left(z'\right)\right) P_h\left(z',S'\right) - P_h(z,S)\right]$$

where $\overline{\delta}(z)$ is the (possibly aggregate state-contingent) average (across houses) depreciation rate of real estate. The real estate company cannot default and is large enough to work under the law of large numbers.

Given that the technology the rental company operates is constant returns to scale and as long as we can ignore the nonnegativity constraint on H_r, profits have to be zero and thus the rental price is given as a known function of the housing price and bond prices

$$P_\ell(z,S) = P_h(z,S) - \sum_{z'} P_b\left(z,S;z'\right)\left(1-\overline{\delta}\left(z'\right)\right) P_h\left(z',S'\right)$$

Note that if the house price P_h and bond prices were constant with $P_b(z,S;z') = \frac{\pi(z'|z)}{1+r}$ then

$$P_\ell = P_h - \frac{(1-\overline{\delta})P_h}{1+r} = \frac{r+\overline{\delta}}{1+r} P_h$$

and the rental price equals the user cost of housing. This expression also pins down the rental price in the steady-state version of our model.

Households

Preferences and endowments

There is a unit mass of households that value nondurable consumption c and housing services h according to the period utility function $u(c,h)$ and discount the future at

rate $\beta \cdot \pi_s$, where β is the discount factor and π_s the survival probability to the next period. Households are assumed to maximize expected utility. We normalize the price of nondurable consumption to 1 in every period.

Specifically, we make the following assumption on preferences. First, the period utility function is given as a CES aggregator over nondurable consumption expenditures c and housing services h,

$$u(c,h) = \frac{\left[(1-\zeta)c^{\nu} + \zeta h^{\nu}\right]^{\frac{1-\sigma}{\nu}}}{1-\sigma} \tag{4.3}$$

where ν governs the elasticity of substitution between consumption and housing.

Second, households can either rent or own houses, but we assume that these two markets are segmented. Houses of size smaller or equal to $\overline{h}$ are only for rent, and houses larger than $\overline{h}$ are only for owner-occupiers.

At the beginning of each period, households choose whether to own or rent, and we denote by V the upper envelope over the renter's value function V_r and the value of owning V_o:

$$V = \max\left\{V_r + \sigma_\xi \xi_r, V_o + \sigma_\xi \xi_o\right\} \tag{4.4}$$

The value functions V, V_r, and V_o are functions of a household's state vector, which we make explicit below but omit for now. Each choice is associated with additive taste shocks ξ_r and ξ_o, which are assumed to be independent and identical draws from a type-I extreme value distribution with scale parameter 1, and σ_ξ is a parameter that governs the importance of taste shocks relative to V_r and V_o. The reason for introducing the taste shocks is two-fold: first, in the Survey of Consumer Finances we observe renter and owner households at each point of the wealth distribution, even though the fraction of home owners approaches 100% among the richest households, which we will be able to capture in our model. In the absence of taste shocks, the model would predict a single cut-off wealth level (conditional on all other states) such that renting would be optional to the left and owning to right of this threshold. Second, as is well know, the expectation over additive type-I extreme-value-distributed shocks gives rise to the closed-form expression

$$\mathbf{E}_{\xi_r,\xi_o}\left[\max\left\{V_r + \sigma_\xi \xi_r, V_o + \sigma_\xi \xi_o\right\}\right] = \sigma_\xi \log\left(e^{V_r/\sigma_\xi} + e^{V_o/\sigma_\xi}\right)$$

which eliminates kinks that would otherwise result from having a discrete choice between renting and owning. Similarly, there is a closed-form expression for the probability of being an owner given by the expression

$$\Pr\left(V_o + \xi_o > V_r + \xi_r\right) = \frac{e^{V_o/\sigma_\xi}}{e^{V_o/\sigma_\xi} + e^{V_r/\sigma_\xi}} = \frac{1}{1 + e^{(V_r - V_o)/\sigma_\xi}} \tag{4.5}$$

which is the well-known probability density function used in bivariate Logit models of discrete choice. Intuitively, the higher the difference $V_r - V_o$, the less likely it is that a household will choose to purchase real estate.

In the remainder of this section we discuss the details of the household problem that are common to both owners and renters. In each period households are endowed with one unit of time that they supply inelastically to the labor market. As is common in the literature, we model the stochastic process for labor productivity y as having a persistent component p and a transitory component ϵ which are additive in logs, so that next-period's y (denoted using primes) is given by

$$\begin{aligned} \log y' &= \log p' + \log \epsilon' + \log \vartheta' \\ \log p' &= \rho_p \log p + \eta' \end{aligned} \tag{4.6}$$

where the innovations are distributed as

$$\begin{aligned} \log \epsilon' &\overset{\text{iid}}{\sim} N\left(-\frac{1}{2}\sigma_\epsilon^2, \sigma_\epsilon^2\right) \\ \log \eta' &\overset{\text{iid}}{\sim} N\left(-\frac{1}{2}\sigma_\eta^2, \sigma_\eta^2\right) . \end{aligned}$$

The choice of mean parameters for both $\log \epsilon$ and $\log \eta$ ensures that the unconditional expectations of p and ϵ in the population are one. We additionally allow for a transitory level shift in household labor productivities represented by ϑ' which only depends on z'. We interpret this as a proxy for unemployment which allows us to vary the aggregate labor supply over the business cycle without having to explicitly introduce an unemployment state.

We apply the Rouwenhorst procedure to discretize p into five states, while ϵ is discretized to three values.[5] Thus in the model labor productivity follows a finite-state Markov chain on a state space of $45 = 5 \times 3 \times 3$ distinct values with transition probabilities $\pi(p', \epsilon', \vartheta' | p, z) = \pi(p'|p)\,\pi(\epsilon')\,\pi(z'|z)$ since $\pi(\vartheta(z')|z) = \pi(z'|z)$. We assume that a law of large numbers applies, so that $\pi(p', \epsilon', \vartheta' | p, z)$ is not only the transition probability for each household, but also the deterministic fraction of the population making a labor productivity transition from (p, z) today to $(p', \epsilon', \vartheta')$ tomorrow. The above assumptions guarantee that aggregate labor supply only depends on the current state z,

$$L(z) = \vartheta(z) \sum_{p,\epsilon} \pi(p)\pi(\epsilon) p\, \epsilon = \vartheta(z) \left[\sum_p \pi(p) p\right]\left[\sum_\epsilon \pi(\epsilon)\epsilon\right] = \vartheta(z)$$

where $\pi(p)$ and $\pi(\epsilon)$ are the cross-sectional fractions of households with persistent and transitory labor productivities p and ϵ, respectively.

5. The transitory shock ϵ does not enter a household's state vector since it is i.i.d. and is only taken into account when computing expectations.

A household's labor income is then given by $w(z, S)y$, where $w(z, S)$ is the economy-wide wage per labor efficiency unit, and households pay labor income taxes according to the (potentially non-linear) tax schedule $T(\bullet)$ so that after-tax earnings are $w(z, S)\,y - T\big(w(z, S)\,y\big)$.

Renters

Conditional on being a renter, a household chooses the optimal level of nondurable consumption c, rental services ρ and savings in Arrow securities $b_{z'}$. The individual state variables are persistent labor productivity p and cash at hand a. Let $s = (a, p)$ denote the individual state variables, and, as before, S as the endogenous aggregate state. The Bellman equation reads

$$\begin{aligned}
V_r\,(s; z, S) &= \max_{c,\rho,b_{z'}} \left\{ u(c,\rho) + \beta\pi_s \mathbf{E}\left[\, V\left(s'; z', S'\right) \Big|\, p, z, S \,\right]\right\} \\
\text{s.t.} \quad a &= c + P_\ell \rho + \sum_{z'} P_b\left(z, S; z'\right) b_{z'} \\
a' &= w\left(z', S'\right) y' - T\big(w\left(z', S'\right) y'\big) + b_{z'} + Tr\left(z', S'\right) \\
y' &= p' \epsilon' \vartheta(z') \\
c &\geq 0\,, \quad h \in \left[0, \overline{h}\right], \quad b_{z'} \geq 0 \;\; \forall\, z' \\
S' &= \Gamma\left(z, S, z'\right)
\end{aligned}$$

where the expectation is taken with respect to z', p', ϵ' and the taste shock realizations (ξ_r', ξ_o'). Tomorrow's cash at hand a' is a function of the household's savings decision as well as shock realizations p', ϵ' and z' at the beginning of next period. Conditional on a tuple (z, z'), the the law of motion for the remaining aggregate state variables in S as perceived by households is given by the transition function $\Gamma(z, S, z')$.

Both renter and owner households can purchase a full set of Arrow securities that pay off contingent on the realization of the aggregate shock z'. Their prices are denoted by $P_b\,(z, S; z')$ and the position chosen by the household as $b_{z'}$.

Households face stochastic survival risk with survival probability π_s, and conditional on survival their continuation value is given by the expression in (4.4). In case of death, a household's savings $b_{z'}$ are collected by the government and returned as part of lump-sump transfers $Tr\,(z, S)$ which are distributed uniformly to all households. In addition to bequests, these transfers include any firm profits distributed to households. We do not impose any "warm-glow" bequest motive.

In the absence of owner-occupied housing, the decision on how to optimally allocate total consumption expenditure between nondurable consumption c and rental services ρ is static. Hence the household problem can be split into a dynamic consumption-savings decision, and for a given level of total consumption $\overline{c}$, an

intratemporal consumption-expenditure problem. For a given rental price $P_\ell(z, S)$, the solution to the latter is denoted as

$$v_r(\bar{c}; P_\ell) \equiv \max_{c,\rho} \left\{ u(c, \rho) \right\} \quad \text{s.t.} \quad \bar{c} = c + P_\ell \rho\,, \quad c \geq 0\,, \quad \rho \in \left[0, \bar{h}\right]$$

The solution to this problem is standard and is relegated to section 4.B in the appendix.

Owners

Households become owners by purchasing real estate that carries idiosyncratic house price risk which they can optionally finance by taking out a mortgage. They additionally have access to the full set of Arrow securities just like the renters.

Real estate is perfectly divisible and has a per-unit price $P_h(z, S)$. Tomorrow the value of the house (per unit) is given by $P_h(z', S')(1 - \delta')$, where δ' is an idiosyncratic house price depreciation shock that is drawn from a differentiable cumulative distribution function $F_{z'}(\delta')$ with domain $[\underline{\delta}(z'), 1]$. As with the idiosyncratic labor productivity shocks, an assumed law of large number assures that $F_z(\bullet)$ is also the cross-sectional distribution over house price depreciation shocks. By assumption households cannot sell short either Arrow securities or real estate.

Mortgages and default. In order to finance the purchase of real estate g households can use short-term mortgage debt. We denote by m the amount of mortgage debt incurred today and to be repaid tomorrow, and by $P_m m$ the amount of resources today advanced to the household for the promise to repay m tomorrow. The mortgage pricing function $P_m(a, p, g, m, \boldsymbol{b}; z, S)$ will be determined in equilibrium and might depend on current household characteristics (a, p), the size of the mortgage m and the housing collateral g, the position of financial assets $\boldsymbol{b} = \left(b_{z_\ell}, b_{z_m}, b_{z_h}\right)$, as well as on aggregate conditions summarized by the aggregate state S. The inverse of the price $1/P_m$ then gives the gross real mortgage interest rate.

The legal environment, which we take as given in this paper, allows households to default on their mortgage debt, with the consequence of having their real estate position foreclosed and a share ϕ of their financial assets confiscated. Thus households default on their mortgage debt tomorrow if and only if

$$P_h\left(z', S'\right)\left(1 - \delta'\right) g + b_{z'} - m < (1 - \phi) b_{z'}$$

that is, if and only if, given the size of the mortgage and the collateral (m, g) and the financial asset position $b_{z'}$, the idiosyncratic house price depreciation shock δ' is sufficiently large. Define the threshold depreciation rate δ^* at which the household is indifferent between defaulting or not by

$$P_h\left(z', S'\right)(1 - \delta^*)\, g + b_{z'} - m = (1 - \phi) b_{z'}$$

or

$$\delta^* = 1 - \frac{m - \phi b_{z'}}{P_h(z', S')\, g} = \delta^*\left(m, g, b_{z'}; z', S'\right) \tag{4.7}$$

and thus the default probability in state z' tomorrow is given by

$$1 - F_{z'}\left(\delta^*\left(m, g, b_{z'}; z', S'\right)\right).$$

Note that if $\phi = 0$ then households default on their mortgage debt if and only if their real estate is under water. On the other hand, if $\phi > 0$ households may refrain from defaulting even in that case (as long as they have some financial assets).

The current household characteristics (a, p) only affect the probability of default tomorrow if they affect the distribution $F_{z'}(\delta')$ of house price shocks tomorrow. Under the assumption that the distribution $F_{z'}(\delta')$ does not depend on (a, p), neither will default probabilities and thus mortgage interest rates. We will already exploit this result and write $P_m(g, m, \boldsymbol{b}; z, S)$ in the recursive formulation of the owner household problem which reads as follows:

$$\begin{aligned}
V_o(s; z, S) = &\max_{c,h,\, b_{z'},g,m,h} \left\{ u(c,h) + \beta\pi_s \mathbf{E}\left[V\left(s'; z', S'\right) \middle| \, p, z, S \right] \right\} + \Delta u_o \\
\text{s.t.} \quad & a = c + \sum_{z'} P_b\left(z, S; z'\right) b_{z'} + g P_h(z, S) - m P_m(g, m, \boldsymbol{b}; z, S) \\
& h = g \\
& a' = w\left(z', S'\right) y' - T\left(w\left(z', S'\right) y'\right) + b_{z'} + Tr\left(z', S'\right) \\
& \qquad + \max\left\{ P_h\left(z', S'\right)\left(1 - \delta'\right) g + b_{z'} - m,\ (1 - \phi) b_{z'} \right\} \\
& y' = p' \epsilon' \vartheta(z') \\
& c \geq 0\,, \quad g \geq \overline{h}\,, \quad b_{z'} \geq 0 \ \ \forall\, z' \\
& S' = \Gamma\left(z, S, z'\right)
\end{aligned}$$

where, as before, $\boldsymbol{b} = \left(b_{z_\ell}, b_{z_m}, b_{z_h}\right)$ is the vector of Arrow bonds contingent on each future z'. Compared to the renter's problem, expectations are additionally taken with respect to the housing depreciation shock δ'. As we do not permit owners to either rent additional housing services beyond their owner-occupied real estate, or to sublet their homes, the amount of rental services consumed is equal to the owner's real estate holdings, i.e., $h = g$. However, owner-occupied housing is associated with an additional flow-utility benefit Δu_o. We will exploit this parameter when calibrating the model to match home ownership rates in the data.

Financial Intermediaries

Representative financial intermediaries compete for mortgages loan by loan, as in Chatterjee et al. (2007), and hence each mortgage is priced to yield zero profits in

expectations. For a loan size m against a house of size $g > 0$ to be repaid tomorrow the intermediary advances $P_m m$ resources today and bears administrative costs $r_w m P_m$. To finance the total cost $(1 + r_w)\, mP_m$ the intermediary sells a position of Arrow securities, specified below. Since survival to the next period is stochastic, we first discuss the optimal default decision of surviving households. In the next period, if the household does not default the financial intermediary collects the repayment of the loan m. However, if the household reneges on its promise, the intermediary obtains the house and sells it for $\gamma\,(1 - \delta')\, gP_h\,(z', S')$ where $\gamma < 1$ measures the efficiency of foreclosure technology. The intermediary also extracts $\phi b_{z'}$ financial assets from the household.

Thus expected revenues from a loan with characteristics $\left(m, g, \boldsymbol{b}\right)$ to a household in aggregate state (z', S') tomorrow are given by $m\Psi_s(m, g, b_{z'}; z', S')$ where Ψ_s is the expected revenue per unit of loan m extended conditional on household survival, which is given by

$$\begin{aligned}\Psi_s\left(m, g, b_{z'}; z', S'\right) = {} & F_{z'}\left(\delta^*\left(m, g, b_{z'}; z', S'\right)\right) \\ & + \phi\left(1 - F_{z'}\left(\delta^*\left(m, g, b_{z'}; z', S'\right)\right)\right)\frac{b_{z'}}{m} \\ & + \frac{\gamma P_h\,(z, S)\, g}{m}\int_{\delta^*(m,g,b_{z'};z',S')}^{1}(1 - \delta')\,dF_{z'}(\delta')\end{aligned} \tag{4.8}$$

On the other hand, a non-surviving household cannot choose to optimally default on its mortgage, so in this case we assume that the intermediary is able to recover

$$\Psi_d\left(m, g; z', S'\right) = \min\left\{P_h\left(z', S'\right)\left(1 - \delta'\right)\frac{g}{m}, 1\right\}$$

per unit of m.[6] Overall, the per-unit expected revenue from a mortgage contract $\left(m, g, \boldsymbol{b}\right)$ next period is therefore

$$\Psi\left(m, g, b_{z'}; z', S'\right) = \pi_s\Psi_s\left(m, g, b_{z'}; z', S'\right) + (1 - \pi_s)\Psi_d\left(m, g; z', S'\right)$$

Since the mortgage market is perfectly competitive loan by loan, for all contracts issued in equilibrium it has to be the case that the cost $(1 + r_w)\, mP_m$ for the intermediary today equals the expected revenue tomorrow $m\Psi\,(m, g, b_{z'}; z', S')$, summed over all aggregate states tomorrow and appropriately discounted to today. Therefore, the equilibrium loan pricing function has to satisfy:

$$(1 + r_w)P_m\left(m, g, \boldsymbol{b}; z, S\right) = \sum_{z'} P_b\left(z, S; z'\right)\Psi\left(m, g, b_{z'}; z', S'\right)$$

6. We assume that there is no inefficiency of foreclose since a non-surviving household does not default, and there is no recourse to Arrow bond savings of a deceased household. Given the low probability of death and the absence of a bequest motive, these assumptions do not matter quantitatively.

or

$$P_m\Big(m, g, \boldsymbol{b}; z, S\Big) = \frac{1}{1 + r_w} \sum_{z'} P_b\left(z, S; z'\right) \Psi\left(m, g, b_{z'}; z', S'\right) \qquad (4.9)$$

Thus what a financial intermediary effectively does when issuing a mortgage m against a house g is to sell Arrow securities in the amounts of $m\Psi\left(m, g, b_{z'}; z', S'\right)$ for each z' to households. It then uses the proceeds

$$m \sum_{z'} P_b\left(z, S; z'\right) \Psi\left(m, g, b_{z'}; z', S'\right)$$

to transfer $mP_m\left(m, g, \boldsymbol{b}; z, S\right)$ to the household that takes the mortgage and bears resource cost $r_w mP_m\left(m, g, \boldsymbol{b}; z, S\right)$. Tomorrow the financial intermediary takes the state-contingent revenues from the mortgages $m\Psi\left(m, g, b_{z'}; z', S'\right)$ to repay its position of the Arrow securities. Thus the mortgage issuer fully hedges against the aggregate risk; by taking on a positive measure of $\left(m, g, \boldsymbol{b}\right)$-type mortgages the law of large numbers assures that the financial intermediary also fully diversifies (by pooling) idiosyncratic household default risk.

Government policy

We model two components of government policy. First, in order to obtain realistic levels of tax rates even in the absence of a government bailout guarantee we assume that the government spends a fixed fraction $\overline{g}$ of total output as nondiscretionary government expenditures:

$$G\left(z, S\right) = \overline{g}Y\left(z, S\right) \qquad (4.10)$$

where $\overline{g} > 0$ is a fixed policy parameter to be calibrated. For future reference we note that

$$G(z, S) = \overline{g}Y(z, S) = \frac{\overline{g}}{1 - \alpha} w(z, S)L(z) \qquad (4.11)$$

Second, and at the core of this paper, we have to model a government bailout guarantee of the financial intermediaries in the real estate sector. We assume that the government bails banks out only when the economy turns from normal times $z \in \left\{z_m, z_h\right\}$ to crisis times $z' = z_\ell$. In that situation, the government guarantees that mortgages, in expectation over idiosyncratic shocks, have the same payoffs for the financial intermediary in good and bad states of the world. That is, the government guarantees that if $z \in \left\{z_m, z_h\right\}$, then

$$\Psi\left(m, g, b_{z_\ell}; z' = z_\ell, S'\right) = \Psi\left(m, g, b_z; z' = z, S'\right)$$

Thus the government insures the mortgage issuer both against lower house prices directly (which affects what the intermediary gets after default) and against the higher

default rates due to lower house prices. The subsidy to a mortgage taker for a mortgage of type $(m, g, \boldsymbol{b})$ is worth today

$$P_b\left(z, S; z_\ell\right)\left[\Psi\left(m, g, b_{z_\ell}; z' = z_\ell, S'\right) - \Psi\left(m, g, b_z; z' = z, S'\right)\right]$$

since the mortgage taker in state $z \in \{z_m, z_h\}$ gets

$$\begin{aligned} P_m\left(m, g, \boldsymbol{b}; z, S\right) &= \frac{1}{1+r_w}\Bigg[P_b\left(z, S; z_\ell\right)\Psi\left(m, g, b_z; z, S'\right) \\ &\qquad + \sum_{z' \neq z_\ell} P_b\left(z, S; z'\right)\Psi\left(m, g, b_{z'}; z', S'\right)\Bigg] \\ &> \frac{1}{1+r_w}\sum_{z'} P_b\left(z, S; z'\right)\Psi\left(m, g, b_{z'}; z', S'\right) \end{aligned} \tag{4.12}$$

In the aggregate state $z' = z_\ell$ tomorrow (conditional on $z \in \{z_m, z_h\}$) the government has to raise taxes to cover the loss to the mortgage issuers, with the total losses equal to

$$T_{FI}\left(z', S'\right) = \int m\left[\Psi\left(m, g, b_{z_\ell}; z' = z_\ell, S'\right) - \Psi\left(m, g, b_z; z' = z, S'\right)\right] d\mu \tag{4.13}$$

where it is understood that m, g and $b_{z'}$ are functions of the state $(s; z, S)$. Note that these required tax revenues tomorrow are a deterministic function of the aggregate states (z, S) today. As a consequence, the parameters governing the tax policy function tomorrow need to be functions of the aggregate state today as well. Rather than spelling this mapping out in full generality we our restrict attention to the class of tax functions used in the quantitative analysis. As in Bénabou (2002) and Heathcote, Storesletten, and Violante (2017) taxes paid on income $\iota = w(z, S)\, y$ are given as

$$T(\iota) = \iota - \lambda \iota^{1-\tau}$$

so that after-tax labor income for a household making per-tax earnings $w(z, S)y$ equals $\lambda\left(w(z, S)\, y\right)^{1-\tau}$. The parameter τ governs the progressivity of the labor income tax schedule and will be treated as fixed in the analysis. The parameter $\lambda = \lambda(z, S)$ instead determines the level of labor income taxes and will vary with the state of the economy to ensure a balanced budget. Total tax revenues are given by

$$\begin{aligned} \int T\left(w(z, S)\, y; z, S\right) d\mu &= \int w(z, S)\, y\, d\mu - \lambda(z, S)\int \left(w(z, S)\, y\right)^{1-\tau} d\mu \\ &= w(z, S)\, L(z) - \lambda w(z, S)^{1-\tau}\int \left(\vartheta(z) p\epsilon\right)^{1-\tau} d\mu \\ &= w(z, S)\, L(z) - \lambda w(z, S)^{1-\tau}\, \vartheta(z)^{1-\tau}\overline{p}_\tau \bar{\epsilon}_\tau \end{aligned}$$

where $\overline{p}_\tau$ and $\overline{\epsilon}_\tau$ are constants defined as

$$\overline{p}_\tau \equiv \sum_p \pi(p)p^{1-\tau} \qquad \text{and} \qquad \overline{\epsilon}_\tau \equiv \sum_\epsilon \pi(\epsilon)\epsilon^{1-\tau}$$

Thus the government budget constraint for tomorrow implies, for $z = \{z_m, z_h\}$ and $z' = z_\ell$,

$$\int T\big(w\left(z', S'\right) y'; z', S'\big) d\mu' = G\left(z', S'\right) + T_{FI}\left(z, S\right)$$

and therefore, using equation (4.10),

$$w\left(z', S'\right) L(z') - \lambda w\left(z', S'\right)^{1-\tau} \vartheta(z')^{1-\tau}\overline{p}_\tau\overline{\epsilon}_\tau = \frac{\overline{g}}{1-\alpha} w\left(z', S'\right) L(z') + T_{FI}\left(z, S\right)$$

For a given level of transfers $T_{FI}\left(z, S\right)$, this can be solved for the tax level in closed form as

$$\lambda(z, S, z') = \frac{w\left(z', S'\right)^\tau}{\vartheta(z')^{1-\tau}\overline{p}_\tau\overline{\epsilon}_\tau}\left[\frac{1-\alpha-\overline{g}}{1-\alpha}L(z') + \frac{T_{FI}\left(z, S\right)}{w\left(z', S'\right)}\right] \tag{4.14}$$

where $S' = \Gamma\left(z, S, z'\right)$ is determined by the aggregate law of motion. Naturally, if there is no bailout tomorrow the second term in brackets is zero.

Recursive competitive equilibrium

We are now ready to define a recursive competitive equilibrium for our economy. The set of endogenous aggregate state variables includes $S = (H, K, \lambda, \mu)$, where H denotes the depreciated housing stock (prior to new construction).[7]

Definition 1. A recursive competitive equilibrium with bailout guarantees are household value and policy functions V_r, V_o, c, h, g, m, $(b_{z'})_{z'}$, aggregate allocation functions I, C_h, I_h, pricing functions P_b, P_m, P_ℓ, P_h, P_k, r, w and an aggregate law of motion Γ such that

1. [Household optimality]: Given prices P_b, P_m, P_ℓ, P_h, w and the aggregate law of motion Γ, the functions V_r and V_o solve the renter and owner household Bellman equations and c, ρ, g, m, $(b_{z'})_{z'}$ are the associated policy functions.

2. [Final goods production firms' optimality]: wages and rental rates of capital are given by

$$w(z, S) = (1-\alpha)A(z)\left(\frac{K(z, S)}{L(z)}\right)^\alpha$$
$$r(z, S) = \alpha A(z)\left(\frac{K(z, S)}{L(z)}\right)^{\alpha-1}$$

7. Since μ is the joint distribution of labor productivity and cash at hand we cannot immediately infer the aggregate capital stock K from it, and thus we include it as an additional state variable. When computing an equilibrium we will use K as one of the state variables anyhow.

3. [Real estate production firms' optimality]: given prices P_h the allocation functions C_h, I_h solve the real estate construction problem (4.2).

4. [Financial intermediaries' optimality]: the mortgage pricing function is given by equation (4.12).

5. Market clearing

   a) The rental market clears

$$H_r = \int \rho\,(s; z, S)\, d\mu$$

   b) The real estate market clears:

$$H_r + \int g\,(s; z, S)\, d\mu = H + I_h\,(z, S)$$

   c) The market for state-contingent assets clears state by state: for all z'

$$\begin{aligned} \int b_{z'}\,(s; z, S)\, d\mu &= B_k\left(z, S; z'\right) \\ &+ \int m\,(s; z, S)\,\Psi\Big(m\,(s; z, S)\,, g\,(s; z, S)\,, \boldsymbol{b}\,(s; z, S)\,; z', S'\Big) d\mu \end{aligned}$$

   d) The goods market clears

$$\begin{aligned} &\int c\,(s; z, S)\, d\mu + I\,(z, S) + \psi_k\left(\frac{I\,(z, S)}{K}\right)K \\ &\quad + C_h\,(z, S) + \psi_h\left(\frac{I_h\,(z, S)}{H}\right)H + \Pi_h\,(z, S) \\ &= A(z)K^{\alpha}L(z)^{1-\alpha} \\ &\quad - r_w \int m\,(s; z, S)\, P_m\Big(m\,(s; z, S)\,, g\,(s; z, S)\,, \boldsymbol{b}\,(s; z, S)\,; z, S\Big) d\mu \end{aligned}$$

   e) No arbitrage between owning physical capital and portfolio of Arrow securities:

$$P_k = \sum_{z'} P_b\left(z, S; z'\right)\left[P_k\left(z', S'\right)(1 - \delta_k) + r\left(s'; z', S'\right)\right]$$

   f) No arbitrage between owning housing and portfolio of Arrow securities:

$$P_\ell\,(z, S) = P_h\,(z, S) + \sum_{z'} P_b\left(z, S; z'\right)\left(1 - \overline{\delta}\left(z'\right)\right) P_h\left(z', S'\right)$$

6. Laws of Motion

   a) Housing stock

$$H' = \Gamma_H(z,S,z') = \int \int_{\underline{\delta}}^{\delta^*(m,g,b_{z'};z',S')} (1-\delta')g(s;z,S)\,dF_{z'}(\delta')\,d\mu$$
$$+\gamma \int \int_{\delta^*(m,g,b_{z'};z',S')}^{1} (1-\delta')g(s;z,S)\,dF_{z'}(\delta')\,d\mu$$

   b) Capital stock

$$K' = \Gamma_K(z,S) = (1-\delta_k)K + I(z,S)$$

   c) Taxes: $\lambda' = \Gamma_\lambda(z,S,z')$ is given by equation (4.14).

   d) Distribution:

$$\mu' = \Gamma_\mu(z,S,z')$$

Remark 1. The corresponding equilibrium without a bailout policy is defined in exactly the same way, but with the mortgage pricing function now being given by (4.9) and the next-period tax rate being

$$\lambda(z,S,z') = \frac{w(z',S')^\tau}{\vartheta(z')^{1-\tau}\overline{p}_\tau\bar{\epsilon}_\tau}\frac{1-\alpha-\overline{g}}{1-\alpha}L(z')$$

4.3 Theoretical characterization

Real estate production sector

The problem of the construction company is, taking as given the aggregate housing stock H, to maximize the revenue from selling newly constructed homes.

$$\begin{aligned} & \max_{I_h,C_h\geq 0} P_h(z,S)I_h - \psi_h\left(\frac{I_h}{H}\right)H - C_h \\ & s.t. \\ I_h &= A_h(z)C_h^{1-\alpha_h} \end{aligned}$$

The first order conditions with respect to I_h, C_h are

$$\begin{aligned} P_h(z,S) - \psi_h'\left(\frac{I_h}{H}\right) &= \lambda \\ \lambda(1-\alpha_h)A_h(z)(C_h)^{-\alpha_h} &= 1 \end{aligned}$$

and thus

$$P_h(z, S) - \psi_h'\left(\frac{I_h}{H}\right) = \frac{(C_h)^{\alpha_h}}{(1-\alpha_h)A_h(z)}$$

and inserting

$$\left(\frac{I_h}{A_h(z)}\right)^{\frac{1}{1-\alpha_h}} = C_h$$

we find

$$P_h(z, S) = \frac{(I_h)^{\frac{\alpha_h}{1-\alpha_h}}}{(1-\alpha_h)A_h(z)^{\frac{1}{1-\alpha_h}}} + \psi_h'\left(\frac{I_h}{H}\right)$$

This equation determines the optimal real estate construction choice I_h as a positive function of the price of real estate, a negative function of the productivity $A_h(z)$ in the real estate sector and a positive function of the existing housing stock (as long as a higher stock negatively affects the marginal cost of adjustment).

Assume that the adjustment cost function is quadratic and defined as

$$\psi_h\left(\frac{I_h}{H}\right) = \frac{1}{2}\overline{\psi}_h\left[\frac{I_h}{H} - \frac{\bar{\delta}(z_m)}{1-\bar{\delta}(z_m)}\right]^2$$

where $\bar{\delta}(z_m)$ is the average depreciation rate in TFP state z_m.[8] Then we have

$$\psi_h'\left(\frac{I_h}{H}\right) = \overline{\psi}_h\left[\frac{I_h}{H} - \frac{\bar{\delta}(z_m)}{1-\bar{\delta}(z_m)}\right]$$

and thus

$$P_h(z, S) = \frac{(I_h)^{\frac{\alpha_h}{1-\alpha_h}}}{(1-\alpha_h)A_h(z)^{\frac{1}{1-\alpha_h}}} + \overline{\psi}_h\left[\frac{I_h}{H} - \frac{\bar{\delta}(z_m)}{1-\bar{\delta}(z_m)}\right]$$

This is an implicit function in I_h that uniquely determines the optimal level of construction as long as $\alpha_h > 0$ or $\psi_h > 0$.

Note that with $\alpha_h = 0$ we have

$$P_h(z, S) = \frac{1}{A_h(z)} + \overline{\psi}_h\left[\frac{I_h}{H} - \frac{\bar{\delta}(z_m)}{1-\bar{\delta}(z_m)}\right]$$

whereas with $\psi_h = 0$ we have

$$P_h(z, S) = \frac{(I_h)^{\frac{\alpha_h}{1-\alpha_h}}}{(1-\alpha_h)A_h(z)^{\frac{1}{1-\alpha_h}}}$$

8. The additional term $\bar{\delta}(z_m)/(1-\bar{\delta}(z_m))$ implies that in a steady-state version of the model with TFP level z_m adjustment costs will be zero.

and with $\psi_h = \alpha_h = 0$ the house price is solely determined by the technology parameter $A_h(z)$,

$$P_h(z, S) = \frac{1}{A_h(z)}$$

Thus as long as there are adjustment costs ($\psi_h > 0$) or curvature in the production function ($\alpha_h > 0$) the supply curve is a positively sloped relation between price P_h and supply of new houses I_h. If $\psi_h = \alpha_h = 0$, the relationship is a horizontal line (if P_h is on the y-axis).

The mortgage interest rate function

Recall that the default threshold is given, for all $g > 0$, by

$$\delta^*\left(m, g, b_{z'}; z', S'\right) = 1 - \frac{m - \phi b_{z'}}{P_h\left(z', S'\right) g}.$$

Defining leverage as $\kappa = m/g$ and the financial-to-real-wealth ratio as $\theta(z') = b_{z'}/g$, we can rewrite this as

$$\delta^* = 1 - \frac{\kappa - \phi\theta(z')}{P_h\left(z', S'\right)} = \delta^*\left(\kappa, \theta(z'); z', S'\right)$$

and thus a household's default probability in aggregate state (z', S') tomorrow only depends on leverage and the financial wealth ratio chosen today.

Conditional on survival, the repayment per unit of loan m in state (z', S') tomorrow thus can be written as

$$\begin{aligned}\Psi_s\left(m, g, b_{z'}; z', S'\right) &= F_{z'}\left(\delta^*\left(\kappa, \theta(z'); z', S'\right)\right) \\ &\quad + \phi\Big(1 - F_{z'}\left(\delta^*\left(\kappa, \theta(z'); z', S'\right)\right)\Big)\frac{\theta(z')}{\kappa} \\ &\quad + \frac{\gamma P_h\left(z, S\right)}{\kappa}\int_{\delta^*\left(\kappa, \theta(z'); z', S'\right)}^{1}(1 - \delta')dF_{z'}(\delta') \\ &= \Psi_s\left(\kappa, \theta(z'); z', S'\right)\end{aligned}$$

and also only depends on leverage and the financial wealth ratio. The same result holds for the per-unit revenue in case of death Ψ_d, which can be written as

$$\Psi_d\left(\kappa; z', S'\right) = \min\left\{\frac{P_h\left(z', S'\right)\left(1 - \delta'\right)}{\kappa}, 1\right\}.$$

This in turn simplifies the mortgage pricing function (in the case without bailout, the expression with bailout is similar) to

$$P_m\left(\kappa, \boldsymbol{\theta}; z, S\right) = \frac{1}{1 + r_w}\sum_{z'} P_b\left(z, S; z'\right)\Psi\left(\kappa, \theta(z'); z', S'\right)$$

where $\boldsymbol{\theta} = (\theta(z'))_{z' \in Z}$ stands in for the entire state-contingent portfolio of financial (relative to real) assets chosen by the household in the current period. Of course, if there is no recourse and $\phi = 0$, then Ψ and consequently P_m is independent of financial asset choices by the household.

We summarize the properties of the mortgage pricing function P_m in the following

Proposition 2. The equilibrium mortgage pricing function has the following properties

1. Mortgages with higher leverage command higher interest rates: $P_m(\kappa, \boldsymbol{\theta}; z, S)$ is strictly decreasing in κ.

2. Households with more financial assets pay lower interest rates: $P_m(\kappa, \boldsymbol{\theta}; z, S)$ is increasing in each $\theta(z')$, strictly so if $\phi > 0$.

Proof. Follows directly from the properties of Ψ. □

4.4 Calibration

Aggregate Risk

We assume that z follows a Markov process at a quarterly frequency defined on the state space (z_ℓ, z_m, z_h) with the transition matrix

$$\Pi_z = \begin{bmatrix} 0.80 & 0.20 & 0.00 \\ 0.05 & 0.85 & 0.10 \\ 0.00 & 0.05 & 0.95 \end{bmatrix}.$$

These probabilities imply that averaged over the business cycle, the economy will be in the depression state z_ℓ approximately 7.7% of the time, in a mild recession in 30.1% of the time and in the expansionary state z_h in the remaining periods.

We allow the aggregate labor supply to vary with z via the shock ϑ in the households' labor productivity process in (4.6). As ϑ is additive in logs and independent of the other determinants of labor productivity, changes in ϑ represent a multiplicative rescaling of labor productivity that is uniform across households. Since we view this as a proxy of unemployment, we calibrate ϑ to correspond to unemployment rates of 12%, 8% and 4% in depressions, mild recession and expansions, respectively. We then normalize ϑ so that it is on average one over the business cycle, resulting in values $(\vartheta_\ell, \vartheta_m, \vartheta_h) = (0.935, 0.977, 1.020)$. Due to the normalizations we apply to persistent and transitory productivity shocks, these figures also correspond to the effective aggregate labor supply $L(z)$ for each z.

	Description	Value	Source
Final-good sector			
α_k	Capital share	0.36	Krueger, Mitman, and Perri (2016)
δ_k	Capital depreciation rate	0.025	Krueger, Mitman, and Perri (2016)
$A(z)$	TFP	0.970, 0.995, 1.005	
$\overline{\psi}_k$	Adj. cost scale	0	
Housing construction			
A_h	Technology constant	0.229	
α_h	Housing constr. elasticity	0.6	Saiz (2010)
$\overline{\psi}_h$	Adj. cost scale	0	
Financial intermediaries			
γ	Foreclosure technology	0.78	Pennington-Cross (2006)
r_w	Mortgage admin. fee	0.001	Jeske, Krueger, and Mitman (2013)

Table 4.1: Parameters for production and financial sectors

Technology

Total factor productivity in the final-goods sector is a function of the aggregate state z given by

$$\big(A(z_\ell), A(z_m), A(z_h)\big) = \big(0.970, 0.995, 1.005\big)$$

We set the capital share to $\alpha = 0.36$ and the capital depreciation rate to $\delta_k = 0.025$ per quarter, as in Krueger, Mitman, and Perri (2016). In the present calibration we assume no capital adjustment costs and set $\overline{\psi}_h = 0$.

For the parameters governing financial intermediation, we closely follow Jeske, Krueger, and Mitman (2013). We set the efficiency of foreclosure to $\gamma = 0.78$ as estimated in Pennington-Cross (2006) and refer to Jeske, Krueger, and Mitman (2013) for a more detailed motivation for this choice. We assume that mortgage servicing costs r_w amount to 10 basis points per period, as in Jeske, Krueger, and Mitman (2013). The technology parameters are summarized in Table 4.1.

Endowments

The parameters of the stochastic endowment process $\left(\rho_p, \sigma_\eta^2, \sigma_\epsilon^2\right)$ are taken from Krueger, Mitman, and Perri (2016), who estimate them from PSID data.

Preferences

Turning to the household preference parameters, as in Jeske, Krueger, and Mitman (2013) we set $\zeta = 0.145$ which is chosen to approximately match the share of housing expenditures in total consumption as reported by the BEA. We set the elasticity of

	Comment	Value
$M/(P_h H_o)$	Aggregate loan-to-value ratio	40%
$P_h H/Y$	Value of housing stock / GDP	6
	Home ownership rate	65%
	Home ownership rate among wealthiest 5%	95%
	P90/P50 house size	2

Table 4.2: Targeted moments of steady-state model (quarterly)

substitution between housing and nondurable consumption to one, which implies that $\nu = 0$ in the CES aggregator in (4.3).

We determine the remaining preference parameters and the minimum house size $\overline{h}$ to match selected moments in *steady-state*.[9] We obtain a steady-state variant of the model by shutting down uncertainty in z, which we set to one. The model is otherwise almost unchanged, except that the three Arrow bonds are replaced with a single risk-free bond.

The target moments are listed in Table 4.2, and selected aggregate variables from the steady-state model are shown in Table 4.3. We attempt to match these by choosing appropriate values of the discount factor β, the relative risk aversion σ, the utility bonus from owning Δu_o, the taste shock scale parameter σ_ξ and the minimum house size $\overline{h}$. The calibrated parameter values are summarized in Table 4.4. We first evaluate the model on a rectangular grid spanned by a few values for each of these parameters, and in a second step use the most promising starting point to minimize the deviation of model and target moments in a weighted least-squares sense.

The strategy for hitting the desired targets is as follows: in order to achieve a realistic aggregate loan-to-value ratio $M/(P_h H_o)$, where M is the aggregate amount of outstanding mortgages and H_o is aggregate owner-occupied housing, we need to set a sufficiently high minimum house size $\overline{h}$ and induce poorer households to take out mortgages to finance their homes. This can be accomplished by increasing Δu_o. The targeted home ownership rate of 65%, however, considerably limits the aggregate mortgage levels this approach can generate since increasing Δu_o at the same time increases home ownership beyond plausible levels. The steady-state model is therefore not able to get all the way to a desired aggregate LTV of 40%.

The other parameter besides Δu_o that strongly affects home ownership rates is the variance of the taste shock controlled by σ_ξ, which we need to balance with the desire to smooth kinks in the household value functions for computational reasons that arise from the discrete choice between renting and owning. Intuitively, taste shocks

9. The full model with aggregate uncertainty takes more than a day to solve on a 28-core machine, and hence it is not feasible to use it for this purpose.

with high variance ensure a higher degree or smoothness, since the importance of the "structural" components V_r and V_o in (4.4) diminishes compared to the innovations (ξ_r, ξ_o). At the same time, a high variance attenuates any differences between the (ex ante) value of being a renter or owner, and the home ownership rate thus approaches 50% across the whole wealth distribution. This is most pronounced for the wealthiest households, since with $\sigma > 1$ the value functions V_r and V_o are bounded from above and consequently V_r and V_o are very close for the rich. We therefore add a home ownership rate of 95% among the wealthiest 5%, which roughly corresponds to the value observed in the Survey of Consumer Finances, as a calibration target to discipline the size of the taste shock.

The relative risk aversion σ predominantly affects the value of the aggregate housing stock relative to GDP, $P_h H/Y$. This ratio is also strongly affected by the utility bonus and the minimum house size $\overline{h}$. We require a value of approximately $\sigma = 6$ to generate $P_h H/Y \approx 5.8$ in steady state.

Lastly, we target a P90/P50 ratio of the house size distribution of two. This target is introduced to prevent too high $\overline{h}$ which would result in all owner households buying the same (minimum) house size, which is clearly counterfactual. This target thus ensures some variation in house sizes observed in our economy.

House depreciation shocks

We adopt the same first two moments for the housing shock as in Jeske, Krueger, and Mitman (2013), who use an annual mean depreciation rate of 1.48% which they obtain from data by the Bureau of Economic Analysis for the years 1960–2002, and an annual standard deviation of 10% based on estimates in Calhoun (1996). For our quarterly calibration, this corresponds to a mean depreciation of about 0.37% and a standard deviation of 5%.

We draw the depreciation shocks from a truncated Frechet distribution with support on $[-0.004, 1.0]$. To this end, we take the probability density function of a non-truncated Frechet distribution with support on $(\underline{\delta}, \infty)$ given by

$$f_{\delta}^{\star}(\delta) = \frac{\alpha_{\delta}}{s_{\delta}} \left(\frac{\delta - \underline{\delta}}{s_{\delta}} \right)^{-1-\alpha_{\delta}} \exp\left\{ - \left(\frac{\delta - \underline{\delta}}{s_{\delta}} \right)^{-\alpha_{\delta}} \right\}$$

and rescale it by the factor $1/F_{\delta}^{\star}(1)$ to obtain the truncated variant $f_{\delta}(\bullet)$, where $F_{\delta}^{\star}(\bullet)$ is the cumulative density function. We use a location parameter $\underline{\delta} = -0.004$ which determines the minimum of the distribution's support, a scale parameter $s_{\delta} = 3.68 \times 10^{-4}$ and a shape parameter $\alpha_{\delta} = 0.6996$. These parameters are chosen to match the first two moments reported above. Compared to the Generalized Pareto distribution used in Jeske, Krueger, and Mitman (2013), the truncated Frechet distribution has the added benefit of being differentiable as $\delta \searrow \underline{\delta}$, which is advantageous as we

	Value	Frac. of GDP
Capital	30.037	8.825
Housing	15.093	–
Risk-free bond price	0.984	–
House price	1.591	–
Rental rate	0.031	–
Aggr. bequests	0.253	0.074
GDP	3.404	1.000
Consumption	2.340	0.688
Gross capital investment	0.751	0.221
Gross housing inv.	0.063	–
Risk-free return	0.016	–
Wages	2.178	–
Rent-price ratio	0.019	–
Owner-occupied housing stock $P_h H_o$	19.802	5.818
Home ownership rate	0.640	–
Share of housing in total cons.	0.167	–
$M/(P_h \times H_o)$	0.308	–
Frac. of owners with mortgage	0.465	–
Median leverage	0.719	–
Default rate (% of owners w/ mort)	0.422	–
Default rate (% of mort. volume)	0.724	–
Housing stock lost due to foreclosure	0.004	–

Table 4.3: Aggregate variables in steady state

	Description	Value	Source
Preferences			
β	Discount factor	0.9797	Endogenously calibrated
π_s	Survival prob.	0.995	Avg. life exp. of 50 years
σ	Risk aversion	6.038	Endogenously calibrated
ν	CES exponent	0	
ζ	CES weight on housing	0.145	Exp. share in BEA
σ_e	Taste shock scale parameter	0.006965	Endogenously calibrated
Δu_o	Owner utility bonus	0.01343	Endogenously calibrated
$\overline{h}$	Min. house size	17.63	Endogenously calibrated
Endowments			
ρ_p	Autocorrelation of persistent shock	0.9923	Krueger, Mitman, and Perri (2016)
σ_η	Conditional std. dev. of persistent shock	0.0991	Krueger, Mitman, and Perri (2016)
σ_ϵ	Std. dev. of transitory shock	0.2285	Krueger, Mitman, and Perri (2016)

Table 4.4: Household preference and endowment parameters

	Description	Value	Source
ϕ	Recourse to financial assets	0.1	
g	Nondiscretionary expenditures (% of GDP)	6%	Brinca et al. (2016)
τ	Tax progressivity	0.181	Heathcote, Storesletten, and Violante (2017)

Table 4.5: Government policy parameters

use a derivatives-based minimizer to solve the household problem. We allow for $F_\delta(\bullet)$ to depend on the aggregate state, but in the current calibration use an identical distribution for all z.

Government policies

We follow Heathcote, Storesletten, and Violante (2017) and set the tax progressivity parameter τ to 0.181, and the share of nondiscretionary government expenditures to 6%. These parameters are summarized in Table 4.5.

4.5 Solution method

In this section we briefly outline the solution method, since it is not entirely standard. For a more detailed discussion we refer to section 4.C in the appendix.

We build on the well-known method of Krusell and Smith (1997, 1998) to handle aggregate uncertainty in a general-equilibrium setting with boundedly-rational agents. Unlike in Krusell and Smith (1998), but similar in spirit to Krusell and Smith (1997), Gomes and Michaelides (2008), and Storesletten, Telmer, and Yaron (2007),

our model needs to ensure market clearing in each simulated period.[10] In contrast to these models, however, we need to clear three to five markets: for housing, up to three Arrow bonds, and capital, if capital adjustment costs are positive and thus the price of capital is no longer unity. To this end, we extend the idea in Storesletten, Telmer, and Yaron (2007) who parametrize the solution to the household problem as a function of the risk premium, and thus they can in each simulation period alter household demand for risk-free bonds and capital by varying prices until market clearing is found.

In our model, we approximate the endogenous aggregate state $S = (K, H, \lambda, \mu)$ by the state vector

$$\widehat{S} = \Big(K, H, P_h, P_\ell, \lambda, Tr, P_b(z_\ell), P_b(z_m), P_b(z_h), P_k\Big)$$

which, together with the law of motion for each of these variables, enables households to solve the renter and owner problems. Given the size of the aggregate state vector, the curse of dimensionality prevents us from solving the household problem on a rectangular grid formed as the Cartesian product of grid points for each of these state variables. We therefore employ a stochastic grid which is sampled from aggregate states observed while simulating the model, as suggested in Maliar, Maliar, and Judd (2011) and Maliar and Maliar (2015).

To illustrate, Figure 4.2 shows scatter plots of the stochastic aggregate grid restricted to two dimensions, the housing stock H and capital K, where the black dots represent grid points conditional on $z = z_m$ today. The colored dots in each panel represent the predicted values tomorrow conditional on $z' = z_\ell$, $z' = z_m$ and $z' = z_h$ in the left-hand, center and right-hand panels, respectively. As the figure shows, agents predict both capital and the housing stock to change only mildly next period.

In contrast, house prices are predicted to have more pronounced movements over the cycle, as depicted in Figure 4.3. House prices drop conditional on being in a depression tomorrow (left-hand panel) and rise in expansions (right-hand panel).

We use global projection methods to evaluate a household's value and policy functions at arbitrary aggregate state vectors that need not be part of the original grid on which the household problem is solved. In the market-clearing step of each simulation period, we are thus able to evaluate each household's policy functions for a candidate price vector $\big(P_h, P_\ell, P_b(z_\ell), P_b(z_m), P_b(z_h), P_k\big)$ and use them to compute aggregate excess demand for each market that needs to be cleared. We vary the price vector until all excess demands are sufficiently close to zero.

10. Market clearing is not required in Krusell and Smith (1998) since the only market that needs to be be cleared is the market for capital (the final-goods market is taken care of by Walras' Law). However, finding a fixed-point of the law of motion for capital in itself guarantees that the capital market is cleared.

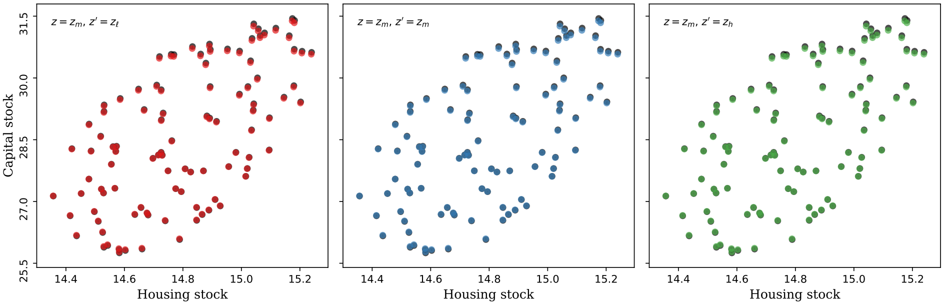


Figure 4.2: Stochastic grid for housing stock H and capital K. Black dots represent grid points for $z = z_m$. Each panel depicts the law-of-motion for a different z' (red dots for $z' = z_\ell$, blue dots for $z' = z_m$ and green dots for $z' = z_h$.

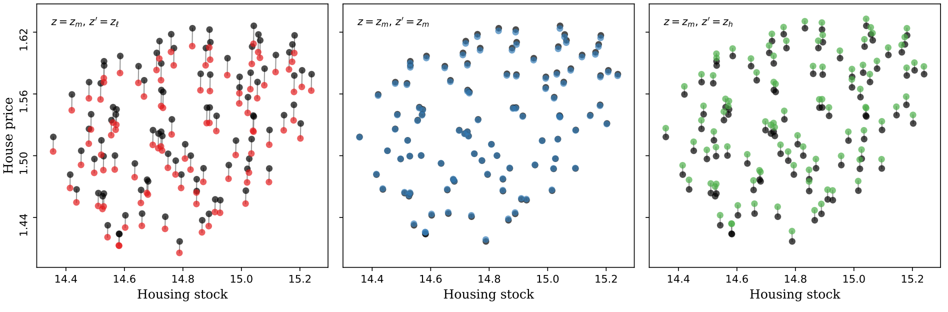


Figure 4.3: Stochastic grid for housing stock H and house price P_h. Black dots represent grid points for $z = z_m$. Each panel depicts the law-of-motion for a different z' (red dots for $z' = z_\ell$, blue dots for $z' = z_m$ and green dots for $z' = z_h$.

	Mean	Std	P01	P25	P50	P75	P99
z: TFP	0.998	0.011	0.970	0.995	1.005	1.005	1.005
Y: GDP	3.376	0.126	3.037	3.309	3.413	3.480	3.521
K: Capital stock	29.617	1.320	25.712	28.862	29.870	30.636	31.437
L: Aggr. labor supply	0.998	0.029	0.935	0.977	1.020	1.020	1.020
H: Housing stock	14.938	0.174	14.386	14.844	14.965	15.057	15.216
Consumption	2.318	0.053	2.196	2.278	2.324	2.361	2.405
Gross capital inv.	0.740	0.071	0.545	0.710	0.783	0.791	0.815
Gross housing inv.	0.063	0.002	0.056	0.062	0.063	0.065	0.065
K/Y	8.774	0.301	7.984	8.604	8.782	8.917	9.548
Risk-free rate	0.016	0.001	0.013	0.015	0.016	0.017	0.020
Return on capital	0.016	0.001	0.013	0.015	0.016	0.017	0.020
Price of capital	1.000	0.000	1.000	1.000	1.000	1.000	1.000
Wages	2.165	0.036	2.056	2.145	2.172	2.192	2.214
P_h: House price	1.576	0.056	1.410	1.546	1.589	1.619	1.645
P_ℓ: Rental rate	0.031	0.001	0.029	0.031	0.031	0.032	0.033
Rent-price ratio	0.020	0.001	0.019	0.020	0.020	0.020	0.022
$P_h H_o/Y$	5.746	0.161	5.284	5.667	5.759	5.822	6.136
Home ownership rate	0.636	0.008	0.609	0.632	0.637	0.642	0.647
$M/(P_h \times H_o)$	0.361	0.019	0.328	0.339	0.367	0.376	0.392
Default rate (% of owners)	0.369	0.018	0.323	0.358	0.373	0.384	0.397
Default rate (% of volume)	0.653	0.019	0.599	0.638	0.653	0.666	0.711

Table 4.6: Summary statistics for simulated aggregate variables (model *with* bailouts)

4.6 Results

Economy with bailouts

We now turn to discussing the results for the model with bailouts, which we view as the benchmark. Thereafter, we contrast some of the findings with those for the model without bailouts.

In Table 4.7, we tabulate summary statistics for simulated aggregate variables in the bailout economy. Compared to the steady-state model (shown in Table 4.3), the average home ownership rate decreases mildly but the aggregate loan-to-value ratio jumps markedly from approximately 30% in the steady-state model to an average value of 36%, which is considerably closer to our calibration target of 40%. This highlights one of the difficulties with using the steady-state model for calibration, as the moments in the model with aggregate uncertainty can turn out to be quite different.

Table 4.7 reports the usual business cycle time series moments for the economy

	$Corr(x_t, x_{t-1})$	$Corr(x_t, Y_t)$	σ_x	σ_x/σ_Y
z: TFP	0.859	0.895	0.011	0.287
Y: GDP	0.930	1.000	0.038	1.000
K: Capital stock	0.999	0.677	0.046	1.203
L: Aggr. labor supply	0.890	0.909	0.029	0.774
H: Housing stock	1.000	0.368	0.012	0.307
Consumption	0.992	0.817	0.023	0.604
Gross capital inv.	0.858	0.898	0.104	2.732
Gross housing inv.	0.951	0.971	0.037	0.967
K/Y	0.909	−0.205	0.034	0.904
Risk-free rate	0.916	0.108	0.080	2.089
Return on capital	0.902	0.228	0.088	2.308
Price of capital	0.000	0.000	0.000	0.000
Wages	0.990	0.676	0.017	0.439
P_h: House price	0.996	0.808	0.036	0.951
P_ℓ: Rental rate	0.984	0.650	0.024	0.627
Rent-price ratio	0.998	−0.542	0.025	0.665
$P_h H_o/Y$	0.925	−0.040	0.028	0.743
Home ownership rate	0.997	0.682	0.013	0.337
$M/(P_h \times H_o)$	0.811	0.340	0.053	1.399
Default rate (% of owners)	0.864	0.853	0.051	1.338
Default rate (% of volume)	0.514	0.239	0.029	0.758

Table 4.7: Moments of simulated (logged) aggregate variables (model *with* bailouts).

with bailouts.[11] In line with standard business cycle models, the volatility of aggregate consumption is substantially lower than that of GDP, while aggregate investment is more volatile.

Turning to housing and mortgages, we find that house prices are more volatile than rental rates, which is in line with data. On the other hand, the model generates pro-cyclical default rates, which we will discuss in greater detail below.

Moving from aggregate to micro behavior, we plot home ownership rates for households at different points of the distribution in Figure 4.4. The graphs should be read as follows: we first compute the average distribution of our Krusell-Smith economy over the entire simulation period (selected percentiles and the Gini coefficient are reported in Table 4.12 in the appendix).[12] We then use the renter and owner value functions

11. Note that unlike in many other papers in the business cycle literature, we do not run the Hodrick-Prescott filter on the simulated time series before computing the moments in Table 4.7. There is no reason to HP filter our data since there is no trend in our model, but this might nevertheless limit the comparability of our moments with those reported elsewhere.

12. Unlike the steady-state model, the Krusell-Smith economy does not have an ergodic distribution over the idiosyncratic states (a, p), but instead depends on the history of aggregate shocks. We therefore use the distribution over (a, p) averaged over time as a proxy.

and the expression in (4.5) to compute the probability that a household will choose to buy a house for a given state vector $\left(s; z, \widehat{S}\right)$. Due to the law of large numbers, this probability coincides with the home ownership rate among households in state $\left(s; z, \widehat{S}\right)$. Figure 4.4 plots these fractions along the cash at hand percentiles of the averaged distribution, disaggregating by persistent labor productivity $(p_1, \ldots, p_5)$ and aggregate state (z_ℓ, z_m, z_h), which are each shown in a separate panel. Moreover, since there is variation in the endogenous aggregate state $\widehat{S}$ even after conditioning on (s, z), the shaded areas represent the range of values observed across the aggregate grid on which the model is solved, while the thicker line presents the median value.

Figure 4.4 thus shows that home ownership is increasing along the cash at hand distribution, and that, all else equal, it is the income-poor households who are the least likely to purchase real estate. For example, the fraction of homeowners among those with the lowest labor productivity p_1 at the median cash at hand level is only about 25%, while it is approximately 90% among the income richest.

For very high wealth levels the home ownership rate approaches a constant level that is independent of any household states and is solely determined by the variance of the taste shock, as discussed in the calibration section. For our calibration this value turns out to be approximately 90%.

In Figure 4.5, we plot the optimal leverage chosen by households at different points of the distribution. These graphs should be read in exactly the same way as those we discussed previously. Note that the figure shows the actual leverage defined as $m/(P_h \cdot g)$ as opposed to $\kappa = m/g$ which is used to solve the model. For the poorest home owners leverage peaks at around 100% as they would not be able to finance their homes otherwise. It decreases along the cash at hand distribution, except for income-poor household where we observe a second peak. This coincides with a jump in bond savings among these households, so they in part take out larger mortgages to invest in financial assets. The reason is that the option to default provides consumption insurance in bad states of the world, while recourse to financial assets in case of default is limited.[13]

We conclude this section with Figure 4.6, which illustrates the endogenous mortgage pricing mechanism at the core of the model. Poorer households with a higher likelihood of default are charged higher markups over the risk-free rate, which peak at around 2.5% for the poorest borrowers. This spread drops quickly along the cash

13. The visual artifacts in the leverage graphs arise because the numerical minimizer struggles to make progress for some aggregate price vectors. This problem occurs because in terms of utility, alternative portfolio positions (high leverage and high bond savings vs. low leverage and low bond savings) are numerically almost equivalent. It becomes more severe for high levels of relative risk aversion such as the one used in our calibration, as then the value functions for even moderate cash at hand levels are effectively flat. The issue disappears for low levels of risk aversion, or when increasing the spread between mortgage interest rates and bond returns, but then the model cannot generate the desired level of aggregate mortgage volumes.

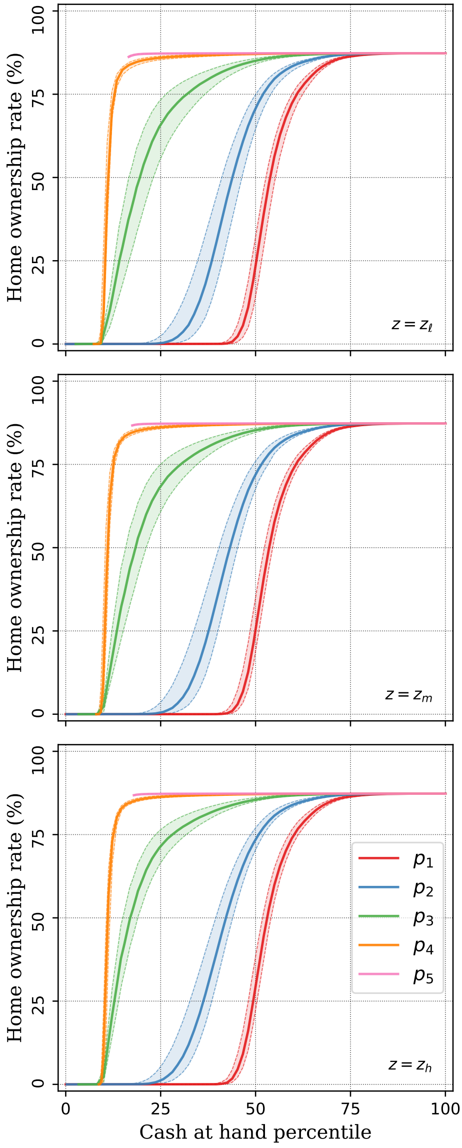


Figure 4.4: Home ownership rates by cash at hand percentile (model *with* bailouts). Plots are disaggregated by persistent labor components $(p_1, \ldots, p_5)$ where p_1 is the lowest productivity level. Each panel shows a different aggregate state $z = (z_\ell, z_m, z_h)$. Shaded areas indicate the variation observed across states $\widehat{S}$ on the aggregate grid.

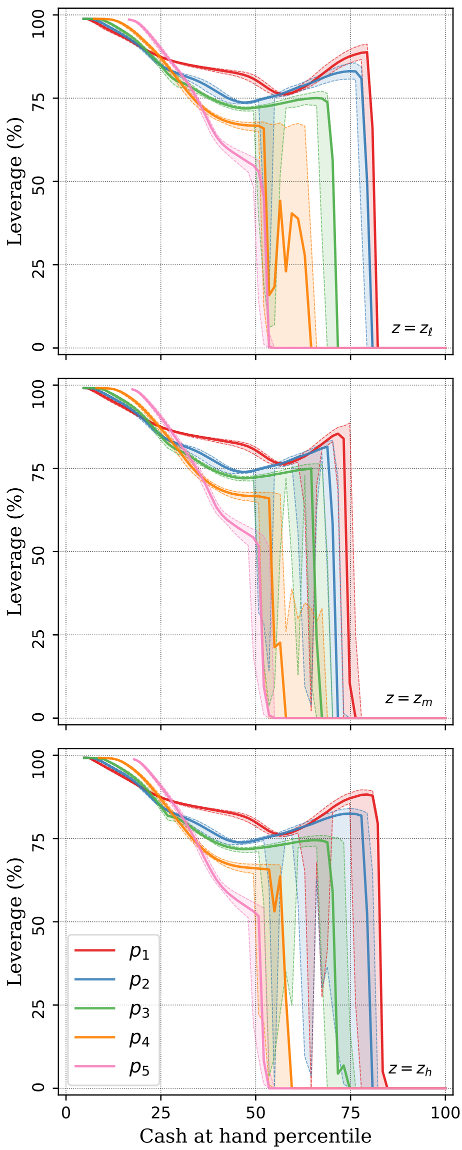


Figure 4.5: Optimal leverage by cash at hand percentile (model *with* bailouts). Plots are disaggregated by persistent labor components $(p_1, \ldots, p_5)$ where p_1 is the lowest productivity level. Each panel shows a different aggregate state $z = (z_\ell, z_m, z_h)$. Shaded areas indicate the variation observed across states $\widehat{S}$ on the aggregate grid.

at hand distribution, amounting to about 25 basis points for the median household. These households have increasing levels of financial wealth so that in the presence of recourse with $\phi > 0$ default becomes less attractive, which is taken into account by financial intermediaries.

Economy without bailouts

The time series summary statistics and moments for the economy without a bailout regime are shown in Table 4.14 and Table 4.15 in the appendix. It turns out that these figures are very similar between the two economies, except for a mild increase in the home ownership rate of two percentage points in the presence of bailouts, and a mild decrease of 0.5 percentage points in the aggregate loan-to-value ratio. These differences are at least partly a result of a shift in the wealth distribution, as households are on average richer in the economy with bailouts (this can be seen by comparing Table 4.12 to Table 4.13 in the appendix).

The rather small aggregate differences arise because quantitatively bailouts are negligible in size, on average amounting to only 0.023% of GDP or 0.011% of outstanding mortgages. While this might be surprising at first, it is a consequence of allowing for partial recourse to financial wealth which *dampens* the households' likelihood to default in the depression state z_ℓ. The reason can be found by looking at the households' financial portfolio choice, and in particular the share invested in Arrow bonds contingent on $z' = z_\ell$, i.e., when a bailout comes into effect (conditional on $z \neq z_\ell$ today). This share is plotted in Figure 4.7. For the poorest households, it approaches 100% of the financial portfolio, since intuitively these households first start saving in Arrow bonds contingent on $z' = z_\ell$ to insure themselves against the particularly bad depression outcome in which they on average experience the most substantial income drop. After the 25th percentile of the cash at hand distribution the share starts dropping towards a value that more or less corresponds to the realization probability of $z' = z_\ell$, which is approximately 80% for $z = z_\ell$ and 5% for $z = z_m$ today.

In our present calibration with recourse to financial wealth in case of default, it may therefore very well happen that household are *less* likely to default in the depression state, precisely because all of their financial savings are invested in the corresponding Arrow bond, while they do not hold Arrow securities contingent on any other outcome. The scope for bailouts is thus greatly diminished. This property of the model helps explain the pro-cyclical aggregate default rates observed in in our simulations.

While the model it its current calibration does not come close to being able to generate bailouts of a magnitude we have observed in the past, it offers some interesting insights at the micro level, which hint at potential distributional effects a plausibly calibrated bailout policy might have. In Figure 4.8, we plot the home ownership rates for both the economies with (solid line) and without (dash-dotted line) bailouts.

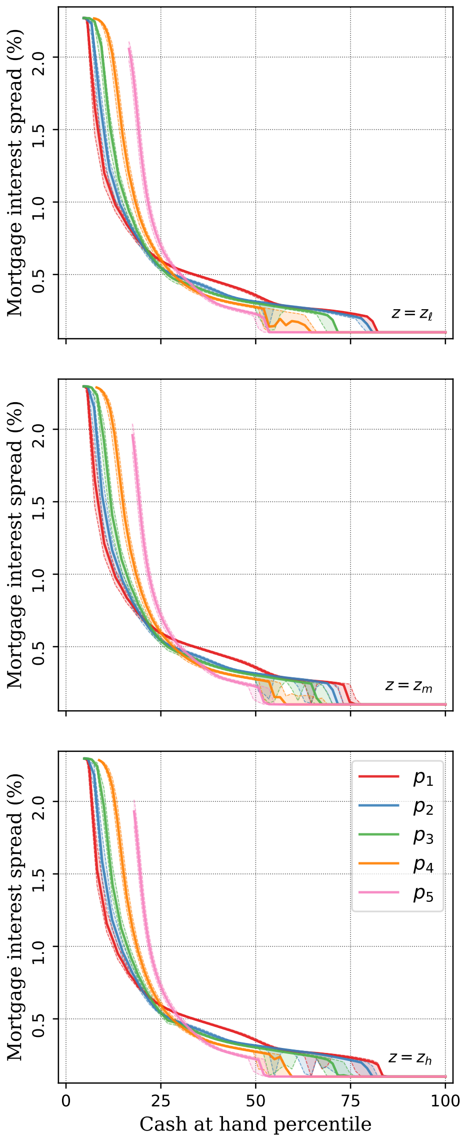


Figure 4.6: Mortgage interest rate spread over the risk-free rate (model *with* bailouts). Plots are disaggregated by persistent labor components $(p_1, \ldots, p_5)$ where p_1 is the lowest productivity level. Each panel shows a different aggregate state $z = (z_\ell, z_m)$. Shaded areas indicate the variation observed across states $\widehat{S}$ on the aggregate grid.

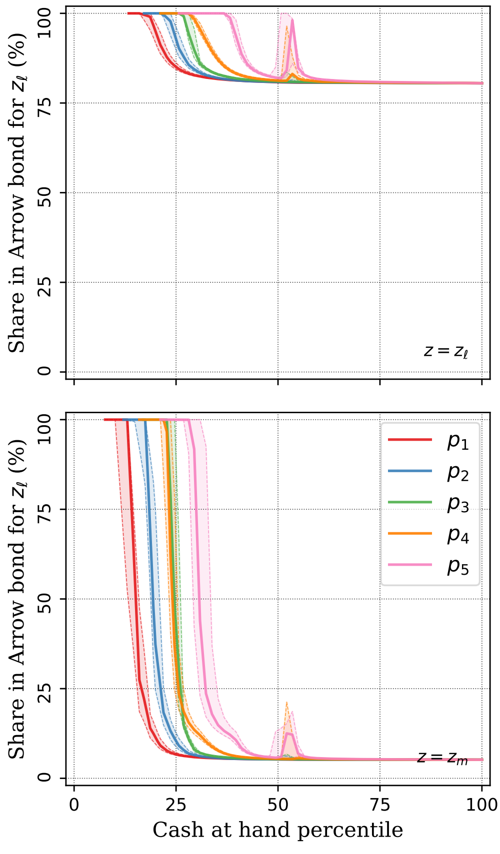


Figure 4.7: Share of financial wealth invested in Arrow bond contingent on $z' = z_\ell$ (model *with* bailouts). Plots are disaggregated by persistent labor components $(p_1, \ldots, p_5)$ where p_1 is the lowest productivity level. Each panel shows a different aggregate state $z = (z_\ell, z_m)$. Shaded areas indicate the variation observed across states $\widehat{S}$ on the aggregate grid.

It is evident that the presence of a bailout policy in particular helps income-poor households to become home owners, while it has only a small effect on the income rich.

Lastly, the presence of bailout guarantees affects on the mortgage interest rate spread set by financial intermediaries. Figure 4.9 contrasts the markups charged across the distribution between the economy with (solid line) and without (dash-dotted line) bailouts.[14] The median mortgage borrower is thus able to profit from a bailout regime, at least before taking into account general-equilibrium effects and a slightly higher average tax rate.

4.7 Conclusion

In this paper, we build a framework which allows us to study government interventions in the housing and mortgage markets in line with recent events such as the Great Recession. In particular, we investigate the effect of government bailout guarantees to financial intermediaries in the mortgage market and their effects on the mortgage interest rates and the allocation of resources between housing and capital used in production.

Our model is able to generate movements in house prices over the business cycle, and the presence of bailout guarantees has effects both on the wealth distribution and on mortgage interest rate spreads, leading to higher home ownership rates especially among poorer households.

While we find that that these guarantees play a role at the micro level, their magnitude in the aggregate is currently too small compared to bailout volumes observed in the Great Recession. We therefore view the present calibration as work in progress, and will for now refrain from making statements about the welfare effects of government bailouts. We plan to address these shortcomings in the next iteration of our paper, for example by allowing the magnitude of housing depreciation shocks to vary with the cycle, which would trigger more frequent defaults when the economy is in a depression, thus increasing the scope for bailout guarantees.

14. We show only three persistent labor income groups and drop the shaded areas as the graphs become too cluttered otherwise.

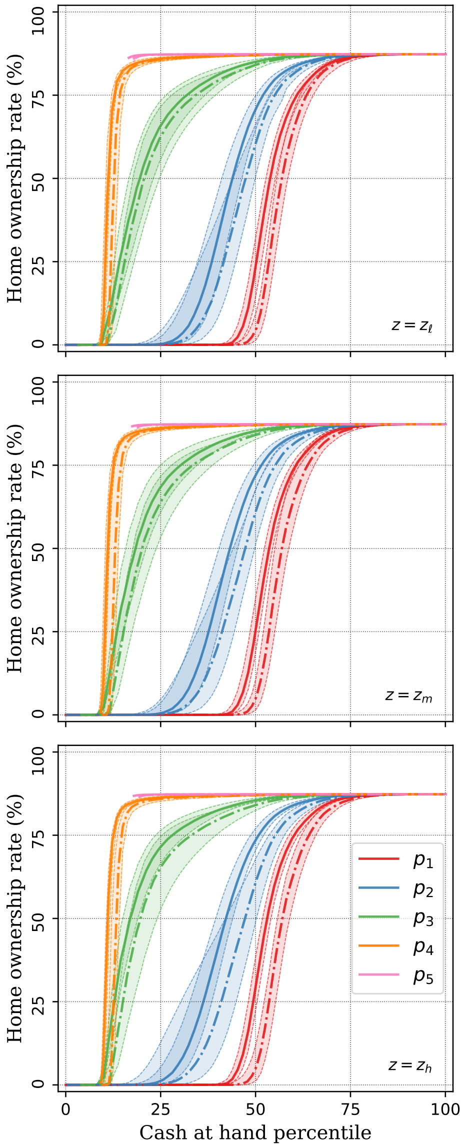


Figure 4.8: Home ownership rates by cash at hand percentile (solid line: model *with* bailouts, dash-dotted line: model *without* bailouts). Plots are disaggregated by persistent labor components $(p_1, \ldots, p_5)$ where p_1 is the lowest productivity level. Each panel shows a different aggregate state $z = (z_\ell, z_m, z_h)$. Shaded areas indicate the variation observed across states $\widehat{S}$ on the aggregate grid.

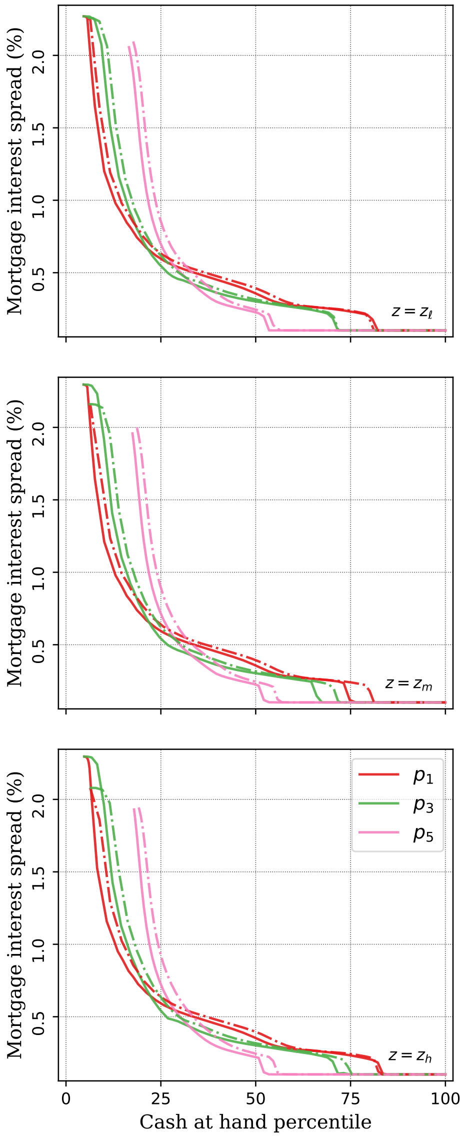


Figure 4.9: Mortgage interest rate spread over the risk-free rate (solid line: model *with* bailouts, dash-dotted line: model *without* bailouts).

Bibliography

Aiyagari, S Rao. 1994. Uninsured idiosyncratic risk and aggregate saving. *The Quarterly Journal of Economics* 109 (3): 659–84.

Bénabou, Roland. 2002. Tax and education policy in a heterogeneous-agent economy: what levels of redistribution maximize growth and efficiency? *Econometrica* 70 (2): 481–517.

Brinca, Pedro, Hans A. Holter, Per Krusell, and Laurence Malafry. 2016. Fiscal multipliers in the 21st century. "Inequality, Institutions, and Redistribution"held at the Stern School of Business,New York University,April 24-25, 2015, *Journal of Monetary Economics* 77:53–69.

Calhoun, Charles A. 1996. Ofheo house price indexes: hpi technical description. *Office of Federal Housing Enterprise Oversight* 20552:1–15.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull. 2007. A quantitative theory of unsecured consumer credit with risk of default. *Econometrica* 75 (6): 1525–1589.

Den Haan, Wouter J. 2010. Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty, *Journal of Economic Dynamics and Control* 34 (1): 79–99.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh. 2016. Phasing out the gses. *Journal of Monetary Economics* 81:111–132.

Favilukis, Jack, Sydney C. Ludvigson, and Stijn Van Nieuwerburgh. 2017. The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy* 125 (1): 140–223.

Frame, W Scott. 2010. *2008 federal intervention to stabilize fannie mae and freddie mac.* DIANE Publishing.

Gervais, Martin. 2002. Housing taxation and capital accumulation. *Journal of Monetary Economics* 49 (7): 1461–1489.

Gomes, Francisco, and Alexander Michaelides. 2008. Asset pricing with limited risk sharing and heterogeneous agents. *Review of Financial Studies* 21 (1): 415–448.

Hayashi, Fumio. 1982. Tobin's marginal q and average q: a neoclassical interpretation [in English]. *Econometrica* 50 (1): 213–224.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. Optimal tax progressivity: an analytical framework. *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Jeske, Karsten, Dirk Krueger, and Kurt Mitman. 2013. Housing, mortgage bailout guarantees and the macro economy. *Journal of Monetary Economics* 60 (8): 917–935.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti. 2017. *Credit supply and the housing boom.* Mimeo, Federal Reserve Banks of Chicago.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante. 2017. *The housing boom and bust: model meets evidence.* Technical report. National Bureau of Economic Research.

Karlman, Markus, Karin Kinnerud, and Kasper Kragh-Sorensen. 2020. *Costly reversals of bad policies: the case of the mortgage interest deduction.* Technical report.

Kraft, Dieter. 1988. *A software package for sequential quadratic programming.* Technical report DFVLR-FB 88-28. Oberpfaffenhofen: Institut für Dynamik der Flugsysteme.

———. 1994. Algorithm 733: tomp–fortran modules for optimal control calculations. *ACM Trans. Math. Softw.* (New York, NY, USA) 20 (3): 262–281.

Krueger, Dirk, Kurt Mitman, and Fabrizzio Perri. 2016. Macroeconomics and household heterogeneity, edited by John B. Taylor and Harald Uhlig, 2:843–921. Handbook of Macroeconomics. Elsevier.

Krusell, Per, and Anthony A. Smith. 1997. Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns. *Macroeconomic Dynamics* 1 (02): 387–422.

———. 1998. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy* 106 (5): 867–896.

Maliar, Lilia, and Serguei Maliar. 2015. Merging simulation and projection approaches to solve high-dimensional problems with an application to a new keynesian model. *Quantitative Economics* 6:1–47.

Maliar, Serguei, Lilia Maliar, and Kenneth Judd. 2011. Solving the multi-country real business cycle model using ergodic set methods. Computational Suite of Models with Heterogeneous Agents II: Multi-Country Real Business Cycle Models, *Journal of Economic Dynamics and Control* 35 (2): 207–228.

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. Household balance sheets, consumption, and the economic slump. *Quarterly Journal of Economics* 128 (4): 1687–1726.

Mian, Atif, and Amir Sufi. 2014. What explains the 2007–2009 drop in employment? *Econometrica* 82 (6): 2197–2223.

Pennington-Cross, Anthony. 2006. The value of foreclosed property. *Journal of Real Estate Research* 28 (2): 193–214.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel. 2007. Housing, consumption and asset pricing. *Journal of Financial Economics* 83 (3): 531–569.

Saiz, Albert. 2010. The geographic determinants of housing supply. *Quarterly Journal of Economics* 125 (3): 1253–1296.

Sommer, Kamila, and Paul Sullivan. 2018. Implications of us tax policy for house prices, rents, and homeownership. *American Economic Review* 108 (2): 241–74.

Storesletten, Kjetil, Chris Telmer, and Amir Yaron. 2007. Asset pricing with idiosyncratic risk and overlapping generations. *Review of Economic Dynamics* 10 (4): 519–548.

Appendix

4.A Investment goods producer

To see that Proposition 1 in the main text holds, denote the Lagrange multiplier on the capital law of motion in the investment firm problem by P_k and conjecture that

$$P_k\left(k; z, S\right) = P_k\left(z, S\right)$$

That is, the Lagrange multiplier, which is Tobin's q or the price of capital, might depend on the aggregate state, but does not depend on the level of capital of the firm. As a consequence, average and marginal P_k are the same, and we can without loss of generality assume the existence of a representative investment firm.

Now let us verify the conjectures, and at the same time characterize the optimal decisions of the firm as well as the associated value functions. Under the conjecture, rewrite the Bellman equation as

$$\kappa_V(z, S) = \max_{\frac{k'}{k}, \frac{i}{k}} \left\{ r\left(z, S\right) - \frac{i}{k} - \psi_k\left(\frac{i}{k}\right) + \frac{k'}{k} \sum_{z'} P_b\left(z, S; z'\right) \kappa_V\left(z', S'\right) \right\}$$

$$\frac{k'}{k} = (1 - \delta_k) + \frac{i}{k}$$

The first order conditions with respect to the choices read as, suppressing dependence on the aggregate state (z, S),

$$P_k = 1 + \psi_k'\left(\frac{i}{k}\right)$$

$$P_k = \sum_{z'} P_b\left(z, S; z'\right) \kappa_V\left(z', S'\right)$$

For a given process of capital returns $r\left(z, S\right)$ and Arrow securities prices $P_b\left(z, S; z'\right)$ these three functional equations (as functions of (z, S)) jointly determine the optimal investment rate $\kappa_i\left(z, S\right)$, the value of the firm per unit of capital $\kappa_V\left(z, S\right)$ and Tobin's q, $P_k\left(z, S\right)$:

$$P_k\left(z, S\right) = 1 + \psi_k'\left(\kappa_i(z, S)\right)$$

$$1 + \psi_k'\left(\kappa_i(z, S)\right) = \sum_{z'} P_b\left(z, S; z'\right) \kappa_V\left(z', S'\right)$$

$$\kappa_V(z, S) = \kappa_d(z, S) + \left[(1 - \delta_k) + \kappa_i(z, S)\right] \sum_{z'} P_b\left(z, S; z'\right) \kappa_V\left(z', S'\right)$$

where the dividend per unit of capital is given by

$$\kappa_d(z,S) = r(z,S) - \kappa_i(z,S) - \psi_k(\kappa_i(z,S))$$
$$d(k;z,S) = \kappa_d(z,S)k$$

4.B Static consumption-expenditure problem

For renter households, the decision of how to allocate total consumption expenditures today is a purely static problem, given the rental price $P_\ell(z,S) > 0$. Thus define

$$v_r(\overline{c};P_\ell) \equiv \max_{c,\rho} \left\{u(c,\rho)\right\} \quad \text{s.t.} \quad \overline{c} = c + P_\ell \rho\,, \quad c \geq 0\,, \quad \rho \in \left[0, \overline{h}\right]$$

Ignoring the constraints on c and ρ for a moment, the solution to this problem is uniquely determined (under the standard concavity, monotonicity and differentiability conditions) by the two equations

$$c + P_\ell \rho = \overline{c}$$
$$\frac{u_\rho(c,\rho)}{u_c(c,\rho)} = \frac{\zeta}{1-\zeta}\frac{\rho^{\nu-1}}{c^{\nu-1}} = P_\ell$$

If the solution to that system of equations implies housing consumption greater than $\overline{h}$, then we set $\rho = \overline{h}$ and back out consumption from the budget constraint. In the case of CES utility given by (4.3), optimal consumption and housing services (for $\rho \leq \overline{h}$) are

$$c = \left[1 - \chi(P_\ell)\right]\overline{c}$$
$$\rho = \chi(P_\ell)\frac{\overline{c}}{P_\ell}$$

where

$$\chi(P_\ell) = \left[1 + P_\ell^{\frac{\nu}{1-\nu}}\left(\frac{1-\zeta}{\zeta}\right)^{\frac{1}{1-\nu}}\right]^{-1}$$

Thus the indirect utility function can be written as

$$v_r(\overline{c};P_\ell) = \frac{\overline{c}^{1-\sigma}\Phi(P_\ell)}{1-\sigma}$$

with

$$\Phi(P_\ell) = \left[\zeta P_\ell^{-\nu}\chi^{\nu-1}\right]^{\frac{1-\sigma}{\nu}}$$

Note that in the special case of a Cobb-Douglas aggregator between consumption and housing services (implied by $\nu = 0$) with consumption share $(1-\zeta)$, we know that the solution will be interior if and only if $\zeta/(1-\zeta) \cdot \underline{c}/P_\ell \leq \overline{h}$. Furthermore,

$$\lim_{\nu \to 0} \chi(P_\ell) = \zeta$$
$$\lim_{\nu \to 0} \Phi(P_\ell) = \left[\zeta^\zeta (1-\zeta)^{1-\zeta} P_\ell^{-\zeta}\right]^{1-\sigma}$$

4.C Numerical solution method

Initial guesses

We generate an initial guess for the stochastic aggregate grid by perturbing the aggregate state of the corresponding steady-state model. Initial guesses for the laws of motion are obtained by log-linearizing the firms' first-order conditions around the steady-state.

Grids

Building on the insight from Krusell and Smith (1998), we approximate the model's true endogenous aggregate state $S = (K, H, \lambda, \mu)$, which contains the distribution of households μ over cash at hand and persistent labor productivity, by the vector $\widehat{S}$ which contains all the continuous aggregate state variables and prices:[15]

$$\widehat{S} = \Big(K, H, P_h, P_\ell, \lambda, Tr, P_b(z_\ell), P_b(z_m), P_b(z_h), P_k\Big)$$

For each discrete state z, we create a stochastic simulated grid of continuous states $\left\{\widehat{S}_{z,i}\right\}_i$ by sampling aggregate state realizations observed while simulating the model. We follow the approach in Maliar, Maliar, and Judd (2011) and Maliar and Maliar (2015) to create a grid that is evenly spaced in all dimensions and covers the "ergodic set" of aggregate states of our model.

Household problem

We solve the household problem using value function iteration (VFI). While the renter problem is a straightforward minimization with three choice variables (the quantities of Arrow bonds to purchase conditional on each z'), there are several complications in the owner's problem that do not allow us to use more sophisticated approaches such as the endogenous-grid method (EGM). First, owners choose to

15. We will from now on refer to prices as "states" when we mean that they are part of the approximate aggregate state vector $\widehat{S}$, even though technically prices are not states in this model.

optimally default at the beginning of next period. Second, owners have a non-linear budget constraint that varies with the chosen leverage and bond savings, since the mortgage price adjusts endogenously. Thirds, the large dimension of the choice set (three Arrow bonds, house size and mortgage level) make a method such as EGM, which takes the choice variables as exogenously given, impractical if not outright infeasible. We therefore solve the household problem using `SLSQP`, a quasi-Newton multi-dimensional minimizer which can handle non-linear constraints (Kraft 1988, 1994).

A particular challenge is to compute expectations with respect to the housing depreciation shock in the owner's continuation value. For standard distributions, this can be done using tabulated quadrature nodes and weights, but no such pre-computed data exist for the truncated Frechet distribution used in this model. An additional complication is that the integration domain changes depending on the default threshold δ^* since we have to compute integrals of the form

$$\int_{\delta^*}^{1} (1-\delta') f_{z'} \left(\delta'\right) d\delta'$$

and quadrature nodes and weights vary with the integration boundaries. To address this problem, we precompute the quadrature data for $N = 9$ quadrature nodes on a fine grid of $\delta^{\star}$, which can be obtained from the first $2 \times N$ moments

$$\int_{\delta^*}^{1} \left(\delta'\right)^n f_{z'} \left(\delta'\right) d\delta' , \qquad n \in \{1, \dots, 2N\}.$$

We do this by first approximating the PDF $f_{z'}$ using cubic piecewise polynomials which are easy to integrate analytically. We verify that our method works by comparing it to results obtained using the adaptive quadrature routines in the `QUADPACK` library (one could of course directly use `QUADPACK`, but this slows down the household problem by a factor of 50–100).

In order to solve the household problem, we need to evaluate the household's value functions at arbitrary points $\widehat{S}' = \Gamma\left(z, \widehat{S}, z'\right)$ which are unlikely to be part of our aggregate grid $\left\{\widehat{S}_{z,i}\right\}_{i,z}$. To this end, conditional on the discrete states (p, z) and a point on the discretized cash at hand grid a, we compute projection coefficients $\widehat{b}_{ij}$ such that

$$V_j\left(\widehat{S}; a, p, z\right) \approx \sum_{i=1}^{k} \widehat{b}_{ij}(a, p, z) \times \varphi_i\left(\widehat{S}\right) + v \qquad j \in \{r, o\}$$

by minimizing the residuals v. To do this, we effectively regress the renter and owner value functions V_r and V_o conditional on some (a, p, z) on the set of basis functions $\{\varphi_i\}_i^k$ which we define as the set of complete polynomials of degree 2 that can be

constructed from the aggregate state vector $\widehat{S}$. More precisely, if the dimension of a vector $\widehat{S}$ is N_s, we define

$$\begin{array}{rlll}
\varphi_0(\widehat{S}) = 1 & & & \\
\varphi_1(\widehat{S}) = K & \varphi_2(\widehat{S}) = H & \dots & \varphi_{N_s}(\widehat{S}) = P_b\left(z, S; z_h\right) \\
\varphi_{N_s+1}(\widehat{S}) = K^2 & \varphi_{N_s+2}(\widehat{S}) = H^2 & \dots & \varphi_{2N_s}(\widehat{S}) = P_b\left(z, S; z_h\right)^2 \\
\varphi_{2N_s+1}(\widehat{S}) = K \cdot H & \varphi_{2N_s+2}(\widehat{S}) = K \cdot P_h & \dots & \\
\vdots & \vdots & \vdots & \vdots
\end{array}$$

Since for our state space this yields several dozen basis functions which are highly collinear, we run principle-component regressions and choose the number of principal components to explain 99.9% of the variance of right-hand-side variables, which usually corresponds to about five principal components.

Having estimated the $\widehat{b}_{ij}$ coefficients, we can then compute value functions for an arbitrary $\widehat{S'}$ conditional on some (a, p, z) as

$$V_j\left(\widehat{S'}; a, p, z\right) = \sum_{i=1}^{k} \widehat{b}_{ij}(a, p, z) \times \varphi_i\left(\widehat{S'}\right) \tag{4.15}$$

Simulation and market clearing

We simulate the model for 40,500 periods and discard the first 500 periods as a "burn-in" phase. The comparatively high number of periods is necessary given our transition matrix for z, as some transitions are observed infrequently even after oversampling them slightly when drawing the sequence of shocks. For example, the transition between from z_m to z_ℓ, which is when bailouts become effective, is observed only approximately 900 times.

In each simulation period we clear markets by finding a price vector

$$\Big(P_h, P_\ell, P_b(z_\ell), P_b(z_m), P_b(z_h), P_k\Big)$$

such that excess demands for housing, all Arrow bonds and capital is close to zero.[16] To this end, we use the multidimensional root-finder provided in the `MINPACK` library. In order to evaluate household policy functions at arbitrary price vectors, we use the exact same projection-method approach as in the household problem, but instead estimate projection coefficients for the household's *policy* functions. We can then determine individual household demand in a way analogous to (4.15), and we sum over the distribution of households to obtain aggregate demand for housing, Arrow

16. For calibrations without capital adjustment costs, we can omit P_k from the candidate price vector.

bonds, etc. After a market-clearing price vector is found, we store the resulting aggregate state vector for each simulation period t,

$$\left(K_t, H_t, P_{h,t}, P_{\ell,t}, \lambda_t, Tr_t, P_{b,t}(z_\ell), P_{b,t}(z_m), P_{b,t}(z_h), P_{k,t}\right)$$

which we use the update the aggregate laws of motion.

Aggregate laws of motion

Using the timeseries of aggregate variables obtained during the simulation, we estimate aggregate laws of motion for each of the state variables in the approximate aggregate state vector $\widehat{S}$ using the following regression equations in the model with bailouts but without capital adjustment costs:

$$\begin{aligned}
\log K' &= \gamma_{1,0}(z) + \gamma_{1,1}(z) \log K + \epsilon_1 \\
\log H' &= \gamma_{2,0}(z, z') + \gamma_{2,1}(z) \log H + \epsilon_2 \\
\log P'_h &= \gamma_{4,0}(z, z') + \gamma_{4,1}(z) \log P_h + \epsilon_4 \\
\log P'_\ell &= \gamma_{5,0}(z, z') + \gamma_{5,1}(z) \log P_\ell + \epsilon_5 \\
\log \lambda' &= \gamma_{6,0}(z, z') + \gamma_{6,1}(z) \log \lambda + \epsilon_6 \\
\log Tr' &= \gamma_{7,0}(z, z') + \gamma_{7,1}(z) \log K + \gamma_{7,2}(z) \log Tr + \epsilon_7 \\
\log P'_{b,z_\ell} &= \gamma_{8,0}(z, z') + \gamma_{8,1}(z, z') \log K + \gamma_{8,2}(z, z') \log H + \epsilon_8 \\
\log P'_{b,z_m} &= \gamma_{9,0}(z, z') + \gamma_{9,1}(z, z') \log K + \gamma_{9,2}(z, z') \log H + \epsilon_9 \\
\log P'_{b,z_h} &= \gamma_{10,0}(z, z') + \gamma_{10,1}(z, z') \log K + \gamma_{10,2}(z, z') \log H + \epsilon_{10} \\
\log T'_{FI} &= \gamma_{12,0}(z, z') + \gamma_{12,1}(z, z') \log K + \gamma_{12,2}(z, z') \log H + \epsilon_{12}
\end{aligned}$$

Note that K is the only predetermined state variable and hence its coefficients only depend on the current-period z, whereas all other variables are jump variables whose coefficients are permitted to be functions of z' or even (z, z').[17] For numerical stability, for some jump variables we nevertheless restrict coefficients to only depend on z since this increases the number of observations these are estimated from.

We use dampened updating to generate a new guess for the coefficients governing the law of motion of each variable. Specifically, if $\widehat{\gamma}_{i,j,n}$ is the estimated coefficient of the j-th r.h.s. variable obtained from running the regressions for the i-th l.h.s. variable in iteration n of the model, we update the coefficient $\gamma_{i,j,n}$ used in the perceived law of motion as a convex combination

$$\gamma_{i,j,n} = (1 - \omega)\gamma_{i,j,n-1} + \omega\widehat{\gamma}_{i,j,n}$$

17. In this model, the housing stock H is not predetermined as the fraction of housing destroyed due to foreclosure depends on z'.

where $\gamma_{i,j,n-1}$ is the coefficient used in iteration $n-1$ and $\omega \in (0,1)$ is the update weight.

An equilibrium in a Krusell-Smith model requires finding a fixed-point of the coefficients γ. For the model with bailouts, we tabulate the R^2 of the estimated laws of motion as well as the changes in coefficients γ between consecutive iterations of the model in Table 4.8. The corresponding statistics for the model without bailouts are shown in Table 4.10. While a high R^2 is often interpreted as a measure of high accuracy of the aggregate laws of motion in Krusell-Smith economies, Den Haan (2010) points out that this can be highly misleading. What's more, if the l.h.s. variable is a jump variable such as a price, a pooled regression which only conditions on $z_t = z$ today, but adds fixed effects for each z_{t+1} tomorrow,

$$\log y_{t+1} = \sum_{z_i} \mathbf{1}\left\{z_{t+1} = z_i\right\} \gamma_{0,i} + \sum_{z_i} \mathbf{1}\left\{z_{t+1} = z_i\right\} \gamma_{1,i} \log x_t + \epsilon_t$$

can have a considerably higher R^2 than running separate regressions on a sub-sample conditioning on $(z_t = z,\ z_{t+1} = z_i)$,

$$\log y_{t+1} = \gamma_{0,i} + \gamma_{1,i} \log x_t + \epsilon_t$$

if the fixed effects explain a substantial part of the variation. However, the estimated coefficients and the prediction error of these two approaches are *identical*. This is something to keep in mind when comparing the R^2 of price equations, in particular the Arrow prices, to those of variables such as K and H.

Den Haan (2010) suggest an alternative approach to assess the accuracy of predictions under the perceived law of motion, which involves computing the average $T_{\mathcal{E}}$-period-ahead prediction error for each simulated state vector $\widehat{S}_t$. To this end, we define the average absolute prediction error for the aggregate variable x starting from period t as

$$\mathcal{E}\left(\widehat{S}_t\right) = \frac{1}{T_{\mathcal{E}}} \sum_{i=1}^{T_{\mathcal{E}}} \left|\widehat{\widehat{x}}_{t+i}\left(\widehat{S}_t\right) - \widehat{x}_{t+i}\right|$$

where $\widehat{x}_{t+i} \in \widehat{S}_{t+i}$ is the actual simulated value of variable x in period $t+i$, whereas $\widehat{\widehat{x}}_{t+i}$ is i-period-ahead prediction conditioning only on information in $\widehat{S}_t$, which is obtained by repeatedly applying the law of motion i times. Similarly, the average relative prediction error is defined as

$$\mathcal{E}_r\left(\widehat{S}_t\right) = \frac{1}{T_{\mathcal{E}}} \sum_{i=1}^{T_{\mathcal{E}}} \left|\frac{\widehat{\widehat{x}}_{t+i}\left(\widehat{S}_t\right) - \widehat{x}_{t+i}}{\widehat{x}_{t+i}}\right|$$

We tabulate the distribution of these errors for a forecast horizon of 30 periods evaluated over the entire simulation in Table 4.9.[18] This shows that the relative

18. Note that for the Arrow prices contingent on $z' = z_\ell$ and $z' = z_m$ we omit those observations where Arrow prices are identically zero, since these artificially lower the absolute prediction error, while the relative prediction error is undefined.

prediction errors are very low for almost all variables including the Arrow prices, despite the low R^2 estimated for some of the equations. The only exception is the relative forecast error for the transfers to financial intermediaries T_{FI} in the economy with bailouts, which is not very well predicted. We have experimented with various r.h.s. specifications for this variable in order to address the issue, but have been unsuccessful so far. However, given the tiny magnitude of aggregate bailouts, even large relative errors will not effect the economy much.

For completeness, the forecast errors for the economy without bailouts are shown in Table 4.10 and are quite low for all variables.

	z	z'	R^2	$\lVert\gamma_n - \gamma_{n-1}\rVert_\infty$	$\lVert\widehat{\gamma}_n - \widehat{\gamma}_{n-1}\rVert_\infty$	$\lVert\gamma_n - \widehat{\gamma}_n\rVert_\infty$
K	z_ℓ		0.99921	5.12e-06	1.17e-03	5.07e-04
	z_m		0.99929	2.26e-05	3.12e-03	2.24e-03
	z_h		0.99967	1.08e-04	2.12e-04	1.07e-02
H	z_ℓ	z_ℓ	0.99995	1.39e-06	5.71e-06	1.38e-04
	z_ℓ	z_m	0.99995	1.42e-06	5.82e-06	1.41e-04
	z_m	z_ℓ	0.99995	2.67e-06	4.60e-05	2.64e-04
	z_m	z_m	0.99995	2.60e-06	4.82e-05	2.57e-04
	z_m	z_h	0.99995	2.56e-06	4.62e-05	2.53e-04
	z_h	z_m	0.99997	1.59e-07	5.59e-06	1.57e-05
	z_h	z_h	0.99997	2.16e-07	4.67e-06	2.14e-05
P_h	z_ℓ	z_ℓ	0.99940	2.13e-05	2.71e-04	2.11e-03
	z_ℓ	z_m	0.99940	2.13e-05	2.71e-04	2.11e-03
	z_m	z_ℓ	0.99924	3.41e-05	9.47e-04	3.38e-03
	z_m	z_m	0.99924	1.54e-05	9.47e-04	1.53e-03
	z_m	z_h	0.99924	1.88e-05	9.47e-04	1.86e-03
	z_h	z_m	0.99967	2.77e-05	1.45e-04	2.74e-03
	z_h	z_h	0.99967	2.08e-05	1.38e-04	2.06e-03
P_ℓ	z_ℓ	z_ℓ	0.99533	2.77e-04	1.77e-03	2.74e-02
	z_ℓ	z_m	0.99533	2.69e-04	1.78e-03	2.66e-02
	z_m	z_ℓ	0.99529	1.85e-04	3.86e-03	1.83e-02
	z_m	z_m	0.99529	1.87e-04	3.83e-03	1.85e-02
	z_m	z_h	0.99529	1.71e-04	3.80e-03	1.70e-02
	z_h	z_m	0.99720	2.11e-04	5.86e-04	2.08e-02
	z_h	z_h	0.99720	1.92e-04	6.18e-04	1.91e-02
λ	z_ℓ	z_ℓ	0.99794	1.13e-05	5.47e-04	1.11e-03
	z_ℓ	z_m	0.99794	1.13e-05	5.47e-04	1.11e-03
	z_m	z_ℓ	0.99968	7.89e-06	9.84e-04	7.81e-04
	z_m	z_m	0.99968	7.89e-06	9.84e-04	7.81e-04
	z_m	z_h	0.99968	7.89e-06	9.84e-04	7.81e-04
	z_h	z_m	0.99975	3.27e-05	6.47e-05	3.24e-03
	z_h	z_h	0.99975	3.27e-05	6.47e-05	3.24e-03
Tr	z_ℓ	z_ℓ	0.99975	5.39e-04	2.37e-02	5.33e-02
	z_ℓ	z_m	0.99975	5.47e-04	2.37e-02	5.41e-02
	z_m	z_ℓ	0.99978	1.93e-03	5.14e-03	1.91e-01
	z_m	z_m	0.99978	1.95e-03	5.17e-03	1.93e-01
	z_m	z_h	0.99978	1.96e-03	5.18e-03	1.94e-01
	z_h	z_m	0.99990	4.13e-04	9.52e-03	4.09e-02
	z_h	z_h	0.99990	3.98e-04	9.54e-03	3.94e-02
$P_b(z_\ell)$	z_ℓ	z_ℓ	0.99878	6.06e-05	2.04e-05	6.00e-03
	z_ℓ	z_m	0.58490	2.23e-04	4.14e-04	2.21e-02
	z_m	z_ℓ	0.99927	5.90e-05	1.94e-05	5.84e-03
	z_m	z_m	0.26342	2.72e-04	1.08e-03	2.69e-02
	z_h	z_m	0.29006	3.09e-04	8.09e-05	3.06e-02
$P_b(z_m)$	z_ℓ	z_ℓ	0.98553	2.33e-04	1.99e-04	2.31e-02
	z_ℓ	z_m	0.99336	4.65e-06	3.82e-04	4.60e-04
	z_m	z_ℓ	0.97896	2.18e-04	7.24e-04	2.16e-02
	z_m	z_m	0.99502	1.22e-05	2.00e-04	1.21e-03
	z_m	z_h	0.25935	1.63e-05	1.79e-03	1.62e-03
	z_h	z_m	0.99708	1.03e-05	3.16e-04	1.02e-03
	z_h	z_h	0.41884	2.39e-05	1.31e-03	2.36e-03
$P_b(z_h)$	z_ℓ	z_m	0.99848	6.33e-05	1.12e-03	6.27e-03
	z_m	z_m	0.99754	9.36e-05	1.20e-03	9.26e-03
	z_m	z_h	0.99783	7.78e-06	1.83e-05	7.70e-04
	z_h	z_m	0.99793	9.29e-05	1.07e-03	9.20e-03
	z_h	z_h	0.99798	8.55e-07	1.64e-04	8.47e-05
T_{FI}	z_m	z_ℓ	0.03396	8.47e-06	2.39e-04	8.38e-04

Table 4.8: Convergence statistics for aggregate laws of motion (model *with* bailouts). Column $||\gamma_n - \gamma_{n-1}||$ shows the sup-norm of changes in γ used to predict next-period states between iterations n and $n-1$. Column $||\widehat{\gamma}_n - \widehat{\gamma}_{n-1}||$ lists the sup-norm of changes in *estimated* coefficients.

	Variable	Mean	P01	P25	P50	P75	P99
$\log_{10}$ errors	K	−0.99	−1.58	−1.18	−0.95	−0.78	−0.55
	H	−2.07	−3.05	−2.32	−2.03	−1.80	−1.34
	q	−4.02	−4.75	−4.22	−3.97	−3.79	−3.49
	P_h	−2.23	−2.90	−2.42	−2.21	−2.03	−1.70
	P_ℓ	−3.89	−4.61	−4.11	−3.85	−3.64	−3.29
	λ	−3.55	−4.19	−3.72	−3.52	−3.36	−3.06
	Tr	−2.74	−3.54	−2.83	−2.68	−2.59	−2.38
	$P_b(z_\ell)$	−4.31	−4.77	−4.50	−4.38	−4.14	−3.70
	$P_b(z_m)$	−4.07	−4.38	−4.16	−4.10	−4.01	−3.64
	$P_b(z_h)$	−4.11	−4.62	−4.23	−4.12	−4.00	−3.41
	T_{FI}	−3.77	−3.94	−3.83	−3.76	−3.70	−3.61
$\log_{10}$ rel. errors	K	−2.46	−3.06	−2.65	−2.42	−2.25	−1.99
	H	−3.24	−4.23	−3.49	−3.20	−2.97	−2.50
	q	−4.01	−4.74	−4.21	−3.97	−3.79	−3.48
	P_h	−2.43	−3.11	−2.62	−2.41	−2.23	−1.89
	P_ℓ	−2.38	−3.11	−2.60	−2.34	−2.14	−1.80
	λ	−3.59	−4.23	−3.75	−3.56	−3.40	−3.10
	Tr	−2.23	−3.00	−2.31	−2.17	−2.09	−1.86
	$P_b(z_\ell)$	−3.30	−3.70	−3.41	−3.29	−3.19	−2.98
	$P_b(z_m)$	−3.10	−3.76	−3.22	−3.05	−2.94	−2.81
	$P_b(z_h)$	−3.55	−4.56	−3.76	−3.53	−3.30	−2.91
	T_{FI}	−0.63	−0.75	−0.67	−0.62	−0.58	−0.52

Table 4.9: Average 30-period-ahead forecast errors for aggregate laws of motion (model *with* bailouts).

	z	z'	R^2	$\lVert\gamma_n-\gamma_{n-1}\rVert_\infty$	$\lVert\widehat{\gamma}_n-\widehat{\gamma}_{n-1}\rVert_\infty$	$\lVert\gamma_n-\widehat{\gamma}_n\rVert_\infty$
K	z_ℓ		0.99968	2.86e-05	5.91e-06	1.88e-03
	z_m		0.99980	1.29e-04	9.41e-04	8.50e-03
	z_h		0.99997	3.68e-05	7.15e-05	2.41e-03
H	z_ℓ	z_ℓ	0.99997	2.26e-06	2.75e-05	1.49e-04
	z_ℓ	z_m	0.99997	2.27e-06	2.75e-05	1.49e-04
	z_m	z_ℓ	0.99997	8.98e-06	1.51e-05	5.89e-04
	z_m	z_m	0.99997	8.98e-06	1.51e-05	5.90e-04
	z_m	z_h	0.99997	8.98e-06	1.52e-05	5.90e-04
	z_h	z_m	0.99998	3.33e-06	2.36e-05	2.19e-04
	z_h	z_h	0.99998	3.37e-06	2.37e-05	2.21e-04
P_h	z_ℓ	z_ℓ	0.99889	2.95e-05	1.20e-04	1.94e-03
	z_ℓ	z_m	0.99889	2.95e-05	1.20e-04	1.94e-03
	z_m	z_ℓ	0.99935	1.92e-05	3.18e-04	1.26e-03
	z_m	z_m	0.99935	1.92e-05	3.18e-04	1.26e-03
	z_m	z_h	0.99935	1.92e-05	3.18e-04	1.26e-03
	z_h	z_m	0.99996	1.39e-05	2.00e-05	9.11e-04
	z_h	z_h	0.99996	1.39e-05	1.14e-05	9.11e-04
P_ℓ	z_ℓ	z_ℓ	0.97723	3.04e-04	8.71e-04	2.00e-02
	z_ℓ	z_m	0.97723	3.13e-04	8.49e-04	2.05e-02
	z_m	z_ℓ	0.98899	2.99e-05	5.13e-04	1.96e-03
	z_m	z_m	0.98899	1.68e-05	5.23e-04	1.10e-03
	z_m	z_h	0.98899	5.62e-06	5.07e-04	3.69e-04
	z_h	z_m	0.99858	4.56e-05	3.30e-05	3.00e-03
	z_h	z_h	0.99858	5.93e-05	2.92e-05	3.90e-03
λ	z_ℓ	z_ℓ	0.99985	8.91e-06	9.34e-07	5.85e-04
	z_ℓ	z_m	0.99985	8.91e-06	9.34e-07	5.85e-04
	z_m	z_ℓ	0.99989	3.81e-05	2.80e-04	2.50e-03
	z_m	z_m	0.99989	3.81e-05	2.80e-04	2.50e-03
	z_m	z_h	0.99989	3.81e-05	2.80e-04	2.50e-03
	z_h	z_m	0.99997	1.09e-05	2.10e-05	7.17e-04
	z_h	z_h	0.99997	1.09e-05	2.10e-05	7.17e-04
Tr	z_ℓ	z_ℓ	0.99994	1.47e-05	1.79e-05	9.67e-04
	z_ℓ	z_m	0.99994	1.34e-05	1.98e-05	8.81e-04
	z_m	z_ℓ	0.99991	5.00e-05	3.77e-04	3.28e-03
	z_m	z_m	0.99991	5.19e-05	3.74e-04	3.41e-03
	z_m	z_h	0.99991	5.19e-05	3.60e-04	3.41e-03
	z_h	z_m	0.99998	3.27e-06	9.85e-06	2.15e-04
	z_h	z_h	0.99998	3.27e-06	1.21e-05	2.15e-04
$P_b(z_\ell)$	z_ℓ	z_ℓ	0.99726	8.55e-06	8.99e-06	5.62e-04
	z_ℓ	z_m	0.74130	4.75e-05	1.31e-04	3.12e-03
	z_m	z_ℓ	0.99521	1.11e-05	1.98e-06	7.27e-04
	z_m	z_m	0.78886	5.38e-05	5.34e-05	3.53e-03
	z_h	z_m	0.86288	6.42e-05	7.62e-05	4.22e-03
$P_b(z_m)$	z_ℓ	z_ℓ	0.98571	4.89e-06	4.79e-05	3.21e-04
	z_ℓ	z_m	0.99790	4.75e-06	1.08e-04	3.12e-04
	z_m	z_ℓ	0.98457	2.32e-05	1.22e-04	1.52e-03
	z_m	z_m	0.99846	2.98e-06	1.13e-04	1.96e-04
	z_m	z_h	0.94990	2.65e-04	1.30e-04	1.74e-02
	z_h	z_m	0.99893	1.64e-05	5.59e-05	1.08e-03
	z_h	z_h	0.97788	2.02e-04	1.92e-04	1.33e-02
$P_b(z_h)$	z_ℓ	z_m	0.98215	1.85e-04	2.02e-04	1.21e-02
	z_m	z_m	0.97181	1.87e-04	2.58e-04	1.23e-02
	z_m	z_h	0.99668	3.31e-05	7.95e-05	2.18e-03
	z_h	z_m	0.96504	1.74e-04	2.55e-04	1.14e-02
	z_h	z_h	0.99871	1.52e-05	3.85e-05	9.99e-04

Table 4.10: Convergence statistics for aggregate laws of motion (model *without* bailouts). Column $||\gamma_n - \gamma_{n-1}||$ shows the sup-norm of changes in γ used to predict next-period states between iterations n and $n-1$. Column $||\widehat{\gamma}_n - \widehat{\gamma}_{n-1}||$ lists the sup-norm of changes in *estimated* coefficients.

	Variable	Mean	P01	P25	P50	P75	P99
$\log_{10}$ errors	K	−1.19	−2.04	−1.42	−1.16	−0.94	−0.57
	H	−2.18	−3.07	−2.43	−2.14	−1.91	−1.49
	q	−4.16	−5.06	−4.38	−4.12	−3.90	−3.42
	P_h	−2.57	−3.40	−2.77	−2.54	−2.35	−1.93
	P_ℓ	−3.92	−4.65	−4.15	−3.89	−3.69	−3.21
	λ	−3.80	−4.67	−4.03	−3.77	−3.55	−3.15
	Tr	−3.19	−4.12	−3.42	−3.14	−2.92	−2.59
	$P_b(z_\ell)$	−4.80	−5.93	−5.10	−4.77	−4.48	−3.87
	$P_b(z_m)$	−4.45	−5.03	−4.65	−4.46	−4.28	−3.70
	$P_b(z_h)$	−4.30	−5.04	−4.52	−4.28	−4.08	−3.68
$\log_{10}$ rel. errors	K	−2.65	−3.52	−2.89	−2.63	−2.40	−2.00
	H	−3.35	−4.24	−3.60	−3.30	−3.08	−2.65
	q	−4.15	−5.05	−4.37	−4.12	−3.90	−3.41
	P_h	−2.76	−3.61	−2.96	−2.73	−2.54	−2.10
	P_ℓ	−2.42	−3.16	−2.65	−2.39	−2.19	−1.71
	λ	−3.84	−4.71	−4.07	−3.81	−3.58	−3.18
	Tr	−2.67	−3.62	−2.90	−2.62	−2.39	−2.06
	$P_b(z_\ell)$	−3.83	−4.68	−4.07	−3.80	−3.56	−3.11
	$P_b(z_m)$	−3.57	−4.27	−3.79	−3.59	−3.35	−2.87
	$P_b(z_h)$	−3.87	−4.70	−4.04	−3.83	−3.67	−3.33

Table 4.11: Average 30-period-ahead forecast errors for aggregate laws of motion (model *without* bailouts).

4.D Additional tables and figures

	Labor	Gini	Mean	Percentiles										
				0.0	1.0	5.0	10.0	25.0	50.0	75.0	90.0	95.0	99.0	100.0
Conditional	1	0.80	11.19	0.45	0.65	0.70	0.77	0.87	1.12	4.35	34.98	71.15	133.24	505.45
	2	0.70	20.63	0.70	1.07	1.42	1.65	2.36	4.35	20.43	67.22	99.29	165.50	568.37
	3	0.56	41.34	1.16	2.13	3.46	4.72	8.96	21.39	56.79	106.79	139.16	217.38	616.04
	4	0.41	80.71	2.06	4.35	10.34	16.60	33.69	65.30	111.98	163.29	204.55	288.69	653.32
	5	0.32	147.43	3.78	9.53	26.93	43.83	82.31	136.18	202.04	263.08	302.84	381.57	694.54
All	–	0.60	50.75	0.45	0.80	1.37	2.21	5.95	25.47	75.22	136.18	177.96	273.45	691.92

Table 4.12: Selected cash at hand percentiles and Gini coefficients for the average distribution of the Krusell-Smith economy (model *with* bailouts). The rows labeled “Conditional” tabulate percentiles conditional on a given persistent labor state, whereas the last row shows moments for all households.

	Labor	Gini	Mean	Percentiles										
				0.0	1.0	5.0	10.0	25.0	50.0	75.0	90.0	95.0	99.0	100.0
Conditional	1	0.82	10.46	0.49	0.62	0.70	0.77	0.87	1.03	2.97	28.82	67.86	145.23	623.39
	2	0.73	19.35	0.73	1.07	1.31	1.53	2.13	3.67	14.23	62.18	100.11	180.29	713.05
	3	0.59	39.09	1.21	2.06	3.26	4.35	7.52	16.60	51.69	105.94	147.29	230.68	787.21
	4	0.43	79.72	2.13	4.23	9.73	15.26	30.80	61.57	109.37	172.23	214.77	301.24	798.57
	5	0.34	147.87	3.89	9.34	26.20	41.87	79.43	134.22	203.29	273.45	314.12	390.80	798.57
All	–	0.62	49.32	0.49	0.80	1.26	1.99	5.11	20.43	71.15	138.16	186.18	285.61	798.57

Table 4.13: Selected cash at hand percentiles and Gini coefficients for the average distribution of the Krusell-Smith economy (model *without* bailouts). The rows labeled “Conditional” tabulate percentiles conditional on a given persistent labor state, whereas the last row shows moments for all households.

	Mean	Std	P01	P25	P50	P75	P99
z: TFP	0.998	0.011	0.970	0.995	1.005	1.005	1.005
Y: GDP	3.351	0.134	2.996	3.274	3.382	3.454	3.547
K: Capital stock	29.007	1.608	24.821	27.955	29.113	30.188	32.139
L: Aggr. labor supply	0.998	0.029	0.935	0.977	1.020	1.020	1.020
H: Housing stock	14.696	0.176	14.200	14.593	14.712	14.818	15.024
Consumption	2.310	0.060	2.164	2.269	2.323	2.357	2.406
Gross capital inv.	0.726	0.079	0.517	0.673	0.764	0.786	0.821
Gross housing inv.	0.062	0.002	0.055	0.061	0.063	0.064	0.066
K/Y	8.656	0.333	7.850	8.443	8.658	8.868	9.469
Risk-free rate	0.017	0.002	0.013	0.016	0.017	0.018	0.021
Return on capital	0.017	0.002	0.013	0.016	0.017	0.018	0.021
Price of capital	1.000	0.000	1.000	1.000	1.000	1.000	1.000
Wages	2.149	0.044	2.030	2.122	2.152	2.180	2.230
P_h: House price	1.556	0.061	1.389	1.518	1.564	1.602	1.662
P_ℓ: Rental rate	0.031	0.001	0.030	0.031	0.031	0.032	0.033
Rent-price ratio	0.020	0.001	0.019	0.020	0.020	0.021	0.022
$P_h H_o/Y$	5.575	0.144	5.218	5.488	5.576	5.660	5.959
Home ownership rate	0.617	0.004	0.603	0.614	0.617	0.620	0.624
$M/(P_h \times H_o)$	0.367	0.010	0.342	0.361	0.368	0.373	0.391
Default rate (% of owners)	0.381	0.037	0.309	0.355	0.378	0.407	0.470
Default rate (% of volume)	0.704	0.057	0.598	0.663	0.697	0.739	0.862

Table 4.14: Summary statistics for simulated aggregate variables (model *without* bailouts)

	$Corr(x_t, x_{t-1})$	$Corr(x_t, Y_t)$	σ_x	σ_x/σ_Y
z: TFP	0.859	0.865	0.011	0.267
Y: GDP	0.940	1.000	0.041	1.000
K: Capital stock	0.999	0.729	0.056	1.379
L: Aggr. labor supply	0.890	0.881	0.029	0.722
H: Housing stock	1.000	0.458	0.012	0.295
Consumption	0.969	0.870	0.026	0.637
Gross capital inv.	0.871	0.930	0.117	2.864
Gross housing inv.	0.954	0.985	0.039	0.966
K/Y	0.927	0.005	0.039	0.944
Risk-free rate	0.941	−0.112	0.090	2.196
Return on capital	0.922	0.010	0.097	2.368
Price of capital	0.000	0.000	0.000	0.000
Wages	0.993	0.729	0.020	0.500
P_h: House price	0.995	0.870	0.040	0.977
P_ℓ: Rental rate	0.988	0.453	0.020	0.491
Rent-price ratio	0.999	−0.709	0.036	0.885
$P_h H_o/Y$	0.918	0.041	0.026	0.631
Home ownership rate	0.990	0.839	0.007	0.167
$M/(P_h \times H_o)$	0.902	0.110	0.028	0.675
Default rate (% of owners)	0.877	0.838	0.096	2.345
Default rate (% of volume)	0.845	0.650	0.079	1.945

Table 4.15: Moments of simulated (logged) aggregate variables (model *without* bailouts).

Sammanfattning

Den här avhandlingen består av fyra fristående kapitel med två gemensamma teman: vad gäller forskningsfråga handlar alla kapitel om hushållens sparande- och portföljbeslut och de underliggande orsakerna till dessa beslut. Vad gäller metod har alla kapitel en gemensam grund i modeller med heterogena agenter, vilket har blivit standard i många områden inom makroekonomi och forskningen kring hushållens finansiella beslut. Målsättningen med dessa metoder är att kunna studera de skillnader i ekonomiskt beteende som orsakas av heterogenitet i exempelvis förmögenhet, förväntad livslängd, förväntningar och husägande, och de olika konsekvenser politiska åtgärder har på dessa hushåll. Detta kopplar naturligt till den ökade användningen av mikrodata för att strukturera makromodeller.

I det första kapitlet, *"Experience-based Learning, Stock Market Participation and Portfolio Choice,"*, undersöker jag en möjlig förklaring till de tydliga mönster vi ser i data vad gäller hushållens portföljsammansättning. Data visar att även i utvecklade ekonomier är deltagandet i aktiemarknaden begränsat och i de flesta länder klart under 50%. Vidare ökar deltagandet med förmögenhet, medan andelen investerad i riskfyllda tillgångar bland de som faktiskt investerat är i princip lika, oavsett förmögenhet, vilket är tvärtemot vad de flesta modeller som används inom forskningen om hushållens finansiella beslut skulle förutsäga.

Nya empiriska resultat tyder på att erfarenheter i livet spelar en viktig roll för hushållens investeringsbeslut. Jag införlivar dessa resultat och det faktum att hushållens finansiella portföljer är underdiversifierade i en för övrigt vanlig livscykelmodell och undersöker till vilken grad detta kan förklara de långvariga öppna frågorna i litteraturen kring deltagande på aktiemarknaden och andel av förmögenheten investerad i riskfyllda tillgångar. Jag visar att lärande genom personliga erfarenheter av avkastning skapar en positiv korrelation mellan hushållens position i förmögenhetsdistributionen och deras optimism avseende framtida avkastning. De förmögna ökar därmed deras investeringar i riskfyllda tillgångar, medan deltagande i aktiemarknaden är begränsat bland fattiga hushåll. Jag finner att i en rimligt kalibrerad kvantitativ modell kan denna mekanism förklara ungefär halva gapet mellan graden av deltagande i data och vad en standardmodell skulle förutsäga.

Kapitel 2 och 3 handlar om heterogenitet i hälsa och förväntad livslängd i befolkningen i USA. Ojämlikhet i hälsa är viktigt i sig, men att förstå de underliggande orsakerna är långt utanför vad som kan omfattas av en avhandling i ekonomi. Vad vi gör är att dokumentera hälsoprocessen på ett struktuerat sätt så att resultatet kan

användas i en makroekonomisk modell, och tar ett litet steg på så sätt att vi undersöker hur ojämlikhet i förväntad livslängd påverkar ekonomiska utfall, speciellt sparande. Många studier har identifierat hälsochocker som en av de största riskerna i livet. En negativ hälsochock kan föranleda stora medicinska utgifter, vilket påverkar incitamenten att spara, och kan också påverka inkomsterna direkt. Den förväntade livslängden har en direkt påverkan på den effektiva diskonteringsfaktorn, en mekanism inbäddad i alla livscykelmodeller med osäker livslängd. Vissa argumenterar vidare att en individs hälsostatus har en direkt påverkan på marginalnyttan från konsumtion. För att kvantifiera risken en individ står inför och modellera de val och beslut individen tar är således en korrekt modellerad hälsoprocess avgörande.

I det andra kapitlet, *"Health dynamics and heterogeneous life expectancies,"* samförfattat med Jonna Olsson, tillhandahåller vi förbättrade estimat av åldersberoende hälsotransitioner och sannolikheter för överlevnad för olika subgrupper i den amerikanska populationen. De estimerade årliga transitionsmatriserna kan användas för alla livscykelmodeller där hälsa och överlevnad är av intresse. Resultaten visar på ansenlig heterogenitet i förväntad livsläng i populationen. En 70-årig man i utmärkt hälsa har en sannolikhet på ca 75% att få uppleva sin åttionde födelsedag, medan motsvarande sannolikhet för en man i dålig hälsa är strax under 40%. Det finns även väsentliga skillnader i förväntad livslängd mellan grupper med olika utbildningsnivå. I gruppen med mindre än motsvarande gymnasieutbildning är den förväntade livslängden vid 50 års ålder 75 år, medan genomsnittet för de med motsvarande universitetsutbildning eller mer är 80 år. Det är två faktorer som bidrar till denna skillnad: för det första, vid 50 års ålder är den genomsnittliga hälsan sämre in gruppen med lägre utbildningsnivå. För det andra, även om vi tar skillnaden i hälsostatus i beaktande, så har gruppen med lägre utbildningsnivå sämre hälsoutveckling och överlevnadssannolikhet även efter 50 års ålder. Vi uppskattar att skillnaden i livslängd mellan olika utbildningsgrupper främst kommer från sämre hälsoutveckling efter 50 års ålder.

I det tredje kapitlet, *"Subjective life expectancies, time preferences heterogeneity and wealth inequality,"* även det samförfattat med Jonna Olsson, använder vi resultaten från det föregående kapitlet och ställer oss den naturliga följdfrågan: vilka konsekvenser har heterogenitet i förväntad livslängd på sparkvoter och i slutändan förmögenhetsojämlikhet?

I enlighet med standardantaganden i ekonomisk teori bör en individ i god hälsa spara mer för framtiden, givet att denna person har en högre sannolikhet att leva ett långt liv. Dock är inte en individs konsumtions- och sparandebeslut nödvändigtvis styrda av objektiv statistisk förväntad livslängd, utan snarare av vad individen uppfattar som sin överlevnadssannolikhet. Vi dokumenterar nya fakta gällande systematisk felskattning i dessa uppfattningar: individer med låg överlevnadssannolikhet relativt sin omgivning underskattar sin överlevnadssannolikhet, medan individer med hög

överlevnadssannolikhet överskattar. Denna systematiska felskattning intensifierar heterogeniteten i förväntad livslängd i populationen.

För att uppskatta effekten av heterogenitet i förväntad livslängd, objektiv såväl som subjektiv, på förmögenhetsojämlikhet använder vi oss av en allmän jämviktsmodell med överlappande generationer där individerna inte kan försäkra sig mot risker. Individerna har heterogena överlevnadssannolikheter som beror på deras aktuella hälsostatus, och deras hälsa förändras beroende på chocker. Utöver denna osäkerhet inkluderar vi även persistenta och övergående chocker till arbetsproduktivitet.

Vi visar att en standardmodell av livscykeltyp har kontrafaktiska implikationer när vi introducerar heterogenitet i förväntad livslängd. I en modell utan arv sparar individer med längre förväntad livslängd mer, som förväntat. Detta stämmer överens med data, där individer med bättre hälsa har större förmögenheter. Men som vi redan känner till, så har en sådan modell utan arv kontrafaktiskt låga förmögenheter bland äldre individer.

Därför inkorporerar vi nytta från att lämna arv i individernas nyttofunktion. Effekten av detta är dock också kontraintuitiv och möjligtvis också oväntad: individer i dålig hälsa sparar nu mer än sina hälsosammare motsvarigheter. Anledningen är följande: eftersom individider i dålig hälsa har en högre sannolikhet att dö snart, lägger de större vikt på den potentiella nyttan från att lämna efter sig ett arv, och därmed har de högre incitament att spara.

Vår slutsats är därmed att ingen av standardmodellerna är lämplig för att studera effekten av heterogenitet i förväntad livslängd på sparkvoter och förmögenhetsojämlikhet. Vi diskuterar potentiella förändringar av standardmodellerna och pekar ut riktningar för framtida forskning.

Slutligen, i det fjärde kapitlet, *"On the Redistributive Effects of Government Bailouts in the Mortgage Market"*, samförfattat med Dirk Krueger och Kurt Mitman, studerar vi vilka faktorer som bestämmer hushållens portföljval när de utöver finansiella tillgångar även innefattar husägande och bolån. Vidare undersöker vi hur statliga åtgärder påverkar dessa beslut, i den mån åtgärderna påverkar bolånekostnader och huspriser, och därmed påverkar hur hushållen fördelar sina resurser mellan fastigheter och finansiella tillgångar.

Mer specifikt studerar vi de aggregerade och fördelningsmässiga konsekvenserna av statliga räddningsaktioner på den amerikanska bolånemarknaden. Vi bygger en modell med aggregerad risk i vilken konkurrensutsatta finansiella intermediärer ger ut bolån till hushåll som kan brista i återbetalningen av sin skuld. Sannolikheten för bristande återbetalning är taget i beaktande i prissättningen av bolåneräntan, såvida inte en garanti om statliga räddningaktioner gör utlånarna opåverkade även i sämre ekonomiska tider då utmätningar ökar. Vi använder modellen till att bedöma i vilken grad statliga garantier leder till alltför stora bolån, alltför hög skuldsättningsgrad bland hushållen, samt ökad volatilitet i huspriser i svåra lågkonjunkturer. Vi studerar

också den snedfördelning mellan fastighetskapital och fysiskt kapital som moralisk risk kan leda till när statliga garantier förekommer. Därmed behöver vi en modell med (stora) aggregerade chocker, fastigheter och kapital som två olika tillgångsslag, samt risk för bristande återbetalning och statliga räddningsaktioner som inte är helt oväntade.

Vi betraktar resultaten i detta kapitel som pågående arbete, och avstår därför från att göra uttalanden om välfärdseffekter av sådana statliga garantier. Vi finner icke-negligerbara fördelningsmässiga konsekvenser och priseffekter av statliga garantier på mikronivå, dock är storleksordningen av de statliga garantierna alltför små jämfört med vad som observerades i den stora recessionen i slutet av 2000-talets första decennium.

IIES Monograph Series

1. Michaely, Michael, *The Theory of Commercial Policy: Trade and Protection*, 1973
2. Söderström, Hans Tson, *Studies in the Microdynamics of Production and Productivity Change*, 1974
3. Hamilton, Carl B., *Project Analysis in the Rural Sector with Special Reference to the Evaluation of Labour Cost*, 1974
4. Nyberg, Lars and Staffan Viotti, *A Control Systems Approach to Macroeconomic Theory and Policy in the Open Economy*, 1975
5. Myhrman, Johan, *Monetary Policy in Open Economies*, 1975
6. Krauss, Melvyn, *International Trade and Economic Welfare,* 1975
7. Wihlborg, Clas, *Capital Market Integration and Monetary Policy under Different Exchange Rate Regimes,* 1976
8. Svensson, Lars E.O., *On Competitive Markets and Intertemporal Resources Allocation*, 1976
9. Yeats, Alexander J., *Trade Barriers Facing Developing Countries*, 1978
10. Calmfors, Lars, *Prices, Wages and Employment in the Open Economy*, 1978
11. Kornai, Jànos, *Economics of Shortage, Vols I and II*, 1979
12. Flam, Harry, *Growth, Allocation and Trade in Sweden. An Empirical Application of the Heckscher-Ohlin Theory*, 1981
13. Persson, Torsten, *Studies of Alternative Exchange Rate Systems. An Intertemporal General Equilibrium Approach*, 1982
14. Erzan, Refik, *Turkey's Comparative Advantage, Production and Trade Patterns in Manufactures. An Application of the Factor Proportions Hypothesis with Some Qualifications*, 1983
15. Horn af Rantzien, Henrik, *Imperfect Competition in Models of Wage Formation and International Trade*, 1983
16. Nandakumar, Parameswar, *Macroeconomic Effects of Supply Side Policies and Disturbances in Open Economies*, 1985
17. Sellin, Peter, *Asset Pricing and Portfolio Choice with International Investment Barriers*, 1990
18. Werner, Ingrid, *International Capital Markets: Controls, Taxes and Resources Allocation*, 1990
19. Svedberg, Peter, *Poverty and Undernutrition in Sub-Saharan Africa: Theory, Evidence, Policy*, 1991

20. Nordström, Håkan, *Studies in Trade Policy and Economic Growth*, 1992
21. Hassler, John, Petter Lundvik, Torsten Persson, and Paul Söderlind, *The Swedish Business Cycle: Stylized facts over 130 years,* 1992
22. Lundvik, Petter, *Business Cycles and Growth*, 1992
23. Söderlind, Paul, *Essays in Exchange Rates, Business Cycles and Growth*, 1993
24. Hassler, John A.A., *Effects of Variations in Risk on Demand and Measures of Business Cycle Comovements*, 1994
25. Daltung, Sonja, *Risk, Efficiency, and Regulation of Banks*, 1994
26. Lindberg, Hans, *Exchange Rates: Target Zones, Interventions and Regime Collapses*, 1994
27. Stennek, Johan, *Essays on Information-Processing and Competition*, 1994
28. Jonsson, Gunnar, *Institutions and Incentives in Monetary and Fiscal Policy*, 1995
29. Dahlquist, Magnus, *Essays on the Term Structure of Interest Rates and Monetary Policy*, 1995
30. Svensson, Jakob, *Political Economy and Macroeconomics: On Foreign Aid and Development,* 1996
31. Blix, Mårten, *Rational Expectations and Regime Shifts in Macroeconometrics*, 1997
32. Lagerlöf, Nils-Petter, *Intergenerational Transfers and Altruism*, 1997
33. Klein, Paul, *Papers on the Macroeconomics of Fiscal Policy*, 1997
34. Jonsson, Magnus, *Studies in Business Cycles*, 1997
35. Persson, Lars, *Asset Ownership in Imperfectly Competitive Markets*, 1998
36. Persson, Joakim, *Essays on Economic Growth*, 1998
37. Domeij, David, *Essays on Optimal Taxation and Indeterminacy*, 1998
38. Flodén, Martin, *Essays on Dynamic Macroeconomics*, 1999
39. Tangerås, Thomas, *Essays in Economics and Politics: Regulation, Elections and International Conflict*, 2000
40. Lidbom, Per Pettersson, *Elections, Party Politics and Economic Policy*, 2000
41. Vestin, David, *Essays on Monetary Policy*, 2001
42. Olofsgård, Anders, *Essays on Interregional and International Political Economics*, 2001
43. Johansson, Åsa, *Essays on Macroeconomic Fluctuations and Nominal Wage Rigidity*, 2002
44. Groth, Charlotta, *Topics on Monetary Policy*, 2002
45. Gancia, Gino A., *Essays on Growth, Trade and Inequality*, 2003
46. Harstad, Bård, *Organizing Cooperation: Bargaining, Voting and Control*, 2003
47. Kohlscheen, Emanuel, *Essays on Debts and Constitutions*, 2004
48. Olovsson, Conny, Essays on Dynamic Macroeconomics, 2004
49. Stavlöt, Ulrika, *Essays on Culture and Trade*, 2005
50. Herzing, Mathias, *Essays on Uncertainty and Escape in Trade Agreements*, 2005

51. Bonfiglioli, Alessandra, *Essays on Financial Markets and Macroeconomics*, 2005
52. Pienaar, Natalie, *Economic Applications of Product Quality Regulation in WTO Trade Agreements*, 2005
53. Song, Zheng, *Essays on Dynamic Political Economy*, 2005
54. Eisensee, Thomas, *Essays on Public Finance: Retirement Behavior and Disaster Relief*, 2005
55. Favara, Giovanni, *Credit and Finance in the Macroeconomy*, 2006
56. Björkman, Martina, *Essays on Empirical Development Economics: Education, Health and Gender*, 2006
57. Larsson, Anna, *Real Effects of Monetary Regimes*, 2007
58. Prado, Jr., Jose Mauricio, *Essays on Public Macroeconomic Policy*, 2007
59. Tonin, Mirco, *Essays on Labor Market Structures and Policies*, 2007
60. Queijo von Heideken, Virginia, *Essays on Monetary Policy and Asset Markets*, 2007
61. Finocchiaro, Daria, *Essays on Macroeconomics*, 2007
62. Waisman, Gisela, *Essays on Discrimination and Corruption*, 2008
63. Holte, Martin Bech, *Essays on Incentives and Leadership*, 2008
64. Damsgaard, Erika Färnstrand, *Essays on Technological Choice and Spillovers*, 2008
65. Fredriksson, Anders, *Bureaucracy, Informality and Taxation: Essays in Development Economics and Public Finance*, 2009
66. Folke, Olle, *Parties, Power and Patronage*, 2010
67. Drott, David Yanagizawa, *Information, Markets and Conflict*, 2010
68. Meyersson, Erik, *Religion, Politics, and Development*, 2010
69. Klingelhofer, Jan, *Models of Electoral Competition*, 2010
70. Perrotta, Maria Carmela, *Aid, Education and Development*, 2010
71. Caldara, Dario, *Essays on Empirical Macroeconomic*s, 2011
72. Mueller, Andreas, *Business Cycles, Unemployment and Job Search*, 2011
73. von Below, David, *Essays in Climate and Labour Economics*, 2011
74. Gars, Johan, *Essays on the Microeconomics of Climate Change*, 2012
75. Spiro, Daniel, *Some Aspects of Resource and Behavioral Economics*, 2012
76. Ge, Jinfeng, *Essays on Macroeconomics and Political Economy*, 2012
77. Li, Yinan, Institutions, *Political Cycles and Corruption*, 2012
78. Håkanson, Christina, *Changes in Workplaces and Careers*, 2013
79. Qin, Bei, *Essays on Empirical Development and Political Economics*, 2013
80. Jia, Ruixue, *Essays on the Political Economics of China's Development*, 2013
81. Campa, Pamela, *Media Influence on Pollution and Gender Equality*, 2013
82. Seim, David, *Essays on Public, Political and Labor Economics*, 2013
83. Shifa, Abdulaziz B., *Essays on Growth, Political Economy and Development*, 2013

84. Panetti, Ettore, *Essays on the Economics of Banks and Markets*, 2013
85. Schmitt, Alex, *Beyond Pigou: Climate Change Mitigation, Policy Making and Decisions*, 2014
86. Rogall, Thorsten, *The Economics of Genocide and War*, 2015
87. Baltrunaite, Audinga, *Political Economics of Special Interests and Gender*, 2016
88. Harbo Hansen, Niels-Jakob, *Jobs, Unemployment, and Macroeconomic Transmission*, 2016
89. Stryjan, Miri, *Essays on Development Policy and the Political Economy of Conflict*, 2016
90. Karadja, Mounir, *On the Economics and Politics of Mobility*, 2016
91. Kitamura, Shuhei, *Land, Power and Technology, Essays on Political Economy and Historical Development*, 2016
92. Malmberg, Hannes, *Human Capital in Development Accounting and Other Essays in Economics*, 2017
93. Öberg, Erik, *On Money and Consumption*, 2017
94. Lane, Nathaniel, *States of Development: Essays on the Political Economy of Development in Asia*, 2017
95. Prawitz, Erik, *On the Move: Essays on the Economic and Political Development of Sweden*, 2017
96. Iwanowsky, Mathias, *Essays in Development and Political Economics*, 2018
97. Dehdari, Sirus Håfström, *Radical Right, Identity, and Retaliation*, 2018
98. Harmenberg, Karl, *Essays on Income Risk and Inequality*, 2018
99. Kilström, Matilda, *Households' Responses to Policy in Labor and Credit Markets*, 2018
100. Meriläinen, Jaakko, *Essays in Political Economics*, 2019
101. Sigurdsson, Jósef , *Essays on Labor Supply and Adjustment Frictions*, 2019
102. Cocciolo, Serena, *Participatory Governance and Public Service Provision* , 2019
103. Mitrunen, Matti, *Essays on the Political Economy of Development*, 2019
104. Olsson, Jonna, *Work, wealth, and well-being: Essays in macroeconomics*, 2019
105. Darougheh, Saman, *Search and Mismatch*, 2019
106. van Vlokhoven, Has, *On the Cost of Capital, Profits and the Diffusion of Ideas*, 2020